



Emergence et entropie : une analyse critique des stratégies explicatives émergentistes basées sur le concept d'entropie

Laurent Jodoin

► To cite this version:

Laurent Jodoin. Emergence et entropie : une analyse critique des stratégies explicatives émergentistes basées sur le concept d'entropie. Philosophie. Université Panthéon-Sorbonne - Paris I; Université de Montréal. Faculté de philosophie, 2015. Français. NNT : 2015PA010503 . tel-01228070

HAL Id: tel-01228070

<https://theses.hal.science/tel-01228070>

Submitted on 12 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Émergence et entropie

par

Laurent Jodoin

Thèse effectuée en cotutelle

présentée à la Faculté des études supérieures et postdoctorales de

l'Université de Montréal en vue de l'obtention du grade de Philosophiae Doctor (Ph.D)

et à

l'Université Paris 1 Panthéon-Sorbonne pour l'obtention du grade de Docteur.

Directeurs de thèse :

Anouk Barberousse, professeur des universités à Lille III
Frédéric Bouchard, professeur à l'université de Montréal

Composition du jury :

Olivier Darrigol, François Duchesneau, Philippe Huneman

17 janvier 2015

©Laurent Jodoin, 2015

Émergence et entropie

*Une analyse critique des stratégies explicatives émergentistes
basées sur le concept d'entropie*

par

Laurent Jodoin

RÉSUMÉ

L'entropie est généralement considérée comme une propriété émergente, tandis que l'émergence de certaines structures organisées serait le résultat d'une dissipation d'entropie. Ainsi, l'émergence est parfois présentée comme ce qui expliquerait l'entropie alors que l'entropie expliquerait l'émergence. Tels quels, ces deux énoncés ne peuvent être tous deux vrais. Face à la polysémie déconcertante des concepts d'émergence, d'entropie et d'explication, je soutiens que cet apparent paradoxe peut être résolu formellement ainsi : l'émergence (en un sens *A*) explique (en un sens *B*) l'entropie (en un sens *C*) et l'entropie (en un sens *D*) explique (en un sens *E*) l'émergence (en un sens *F*). La solution revient donc à préciser *A*, *B*, *C*, *D*, *E*, et *F*. Pour ce faire, je propose un modèle pluraliste (restreint) de l'explication et un examen critique du concept d'entropie.

Dans le cas de l'entropie comme explanandum de l'émergence (*A*, *B* et *C*), ce qui émerge est l'irréversibilité comme propriété essentielle de l'entropie thermodynamique, mais l'émergence ne peut être synonyme de non-explicabilité. Je montre alors trois possibilités où l'émergence peut expliquer l'entropie thermodynamique : selon (i) un sens fort, comme une modalité ontologique, (ii) selon un sens intermédiaire d'après ce que j'appelle l'émergence méthodologique (où il y a possibilité d'explication réductive sans réduction dérivationnelle), et (iii) selon un sens faible, comme sa désignation comme membre d'une classe d'émergents.

Dans le cas de l'entropie comme explanans de l'émergence (*D*, *E*, et *F*), il faut distinguer l'approche substantielle de l'approche analogique. Dans le premier cas, l'entropie renvoie à une propriété macroscopique robuste et autonome pouvant être mobilisée au sein d'un explanans de l'émergence de nouvelles structures complexes. Dans le second cas, l'entropie exemplifie la réalisabilité multiple et peut être mobilisée, modulo une justification, au sein d'un explanans de l'émergence de propriétés à des niveaux supérieurs. En définitive, la polysémie de ces concepts peut être fructueuse au sein de ce cadre explicatif de phénomènes divers et complexes, de la physique à la biologie.

Mots-clés : entropie, émergence, explication, complexité, physique statistique, génétique des populations.

ABSTRACT

Entropy is generally considered as an emergent property, while the emergence of certain organized structures is supposed to be the result of entropy dissipation. Thus, emergence is sometimes seen as explaining entropy, whereas sometimes the explanation is the other way around, as entropy would explain emergence. As such both statements cannot be true. Faced with the daunting polysemy of these concepts of emergence, entropy and explanation, I argue that this apparent paradox can be formally solved as follows: the emergence (in a sense *A*) explains (in a sense *B*) entropy (in a sense, *C*) and entropy (in a sense *D*) explains (in a sense *E*) emergence (in a sense *F*). The solution is therefore to specify *A*, *B*, *C*, *D*, *E*, and *F*. To do this, I suggest a (restricted) pluralistic model of explanation and a critical examination of the concept of entropy.

In the case of entropy as explanandum of emergence (*A*, *B* and *C*), what is emerging is irreversibility as an essential property of thermodynamic entropy, but it cannot be emergence of synonymous of non-explicability. I then show three possibilities that can explain the emergence of thermodynamic entropy: (i) in a strong sense, as an ontological modality, (ii) in an intermediate sense, from what I call ‘methodological emergence’ (where there is a possibility of reductive explanation but no derivational reduction), or (iii) in a weak sense, as its designation as a member of an emergence class.

In the case of entropy as explanans of emergence (*D*, *E*, and *F*), one must distinguish the substantive approach from the analogical approach. In the first case, entropy refers to a robust and autonomous macroscopic property that can be mobilized in an explanans of the emergence of new complex structures. In the second case, the entropy exemplifies multiple realizability and can be mobilized, with a proper justification, within an explanans of the emergence of properties at higher levels. Ultimately, the polysemy of these concepts can be fruitful in this explanatory framework for various complex phenomena, from physics to biology.

Keywords: entropy, emergence, explanation, complexity, statistical physics, population genetics.

TABLE DES MATIÈRES

1	INTRODUCTION GÉNÉRALE.....	12
1.1	<i>Philosophie des sciences</i>	15
1.1.1	Mise en contexte	15
1.1.2	Complexité	17
1.2	<i>Les 3 E, la fourmi et le chapeau</i>	20
1.2.1	Entropie	21
1.2.2	Explication	24
1.2.3	Émergence.....	26
1.2.4	Explanandum ou explanans ?.....	27
1.3	<i>Plan de la thèse</i>	30
2	ÉMERGENCE ET EXPLICATION	31
2.1	<i>Explication</i>	36
2.1.1	Orthodoxie : lois et causes	41
2.1.1.1	Explication probabiliste	45
2.1.1.2	Fictions et idéalizations	45
2.1.2	Alternatives	49
2.1.3	Réduction	55
2.1.3.1	Taxonomie réductionniste.....	56
2.1.3.2	Orthodoxie et critiques.....	57
2.1.4	Un modèle d'explication.....	63
2.1.4.1	Modalité	65
2.1.4.2	Naturalité.....	67
2.2	<i>Niveaux d'organisation</i>	69
2.2.1	Un cadre réaliste.....	69
2.2.2	Conceptualisation des niveaux	72
2.2.3	Inobservables.....	75
2.2.4	Survenance	77
2.2.5	Réalisabilité multiple	79
2.3	<i>Émergence</i>	83
2.3.1	Taxonomie(s) de l'émergence.....	85
2.3.2	Conceptualisation de l'émergence	87
2.3.2.1	Relata.....	88
2.3.2.2	Relation	90
2.3.2.3	Causalité descendante	95
2.3.3	Conclusion préliminaire	98
2.4	<i>Explication et émergence</i>	100
2.4.1	Un modèle : la « stratégie R ».....	101
2.4.2	Émergence expliquée	103

2.4.3	Émergence qui explique.....	108
2.5	<i>Conclusion</i>	111
3	HISTORIQUE DU CONCEPT D'ENTROPIE	114
3.1.1	Révolution scientifique	116
3.1.2	Mise en contexte historique du concept d'entropie.....	119
3.1.3	Le concept d'énergie	122
3.2	<i>Thermodynamique</i>	126
3.2.1	Sadi Carnot.....	128
3.2.2	William Thomson, alias Baron Kelvin.....	132
3.2.3	Rudolf Clausius	135
3.2.4	Aparté : Rankine et Poincaré.....	140
3.2.5	Entropie de non-équilibre.....	140
3.3	<i>Mécanique statistique</i>	144
3.3.1	James Clerk Maxwell	146
3.3.2	Ludwig Boltzmann.....	149
3.3.3	Josiah Willard Gibbs	157
3.3.4	Aparté : Planck et Einstein	161
3.4	<i>Autres disciplines</i>	164
3.4.1	Théorie de l'information	164
3.4.2	Mécanique quantique	166
3.4.3	Biologie	167
3.5	<i>Conclusion</i>	170
4	CRITIQUE DU CONCEPT D'ENTROPIE I	172
4.1	<i>Irréversibilité</i>	178
4.1.1	Sens et référence	180
4.1.1.1	Invariance par renversement du temps.....	181
4.1.1.2	Irrécupérabilité	182
4.1.2	Asymétrie temporelle	186
4.1.2.1	La question du temps	187
4.1.2.1.1	Caractérisation du temps	188
4.1.2.1.2	Temps et historicité	193
4.1.2.2	Le temps en physique.....	198
4.1.2.3	Flèche du temps.....	201
4.2	<i>Thermodynamique</i>	208
4.2.1	États et processus en thermodynamique	211
4.2.2	Principes fondamentaux	213
4.2.3	Le second principe de la thermodynamique.....	216
4.2.3.1	Reconstruction.....	218
4.2.3.2	Discussion	220
4.2.3.3	Critiques	228

4.2.4	La fonction d'état S	233
4.2.4.1	Propriétés	238
4.2.5	Principe entropique	240
4.2.5.1	Caractérisation	241
4.2.5.2	Critiques	247
4.2.6	Non-équilibre	253
4.2.7	Stratégie explicative	256
4.2.7.1	Stratégie généralisée	261
4.2.7.2	Discussion : limites et critiques	263
4.2.7.3	Remarques conclusives	268
4.3	<i>Conclusion</i>	270
5	CRITIQUE DU CONCEPT D'ENTROPIE II	273
5.1	<i>Probabilités</i>	276
5.1.1	Introduction	276
5.1.2	Interprétations	279
5.1.3	Discussion	286
5.1.3.1	Schème objectif-subjectif	288
5.1.3.2	Probabilité en physique	294
5.1.4	Remarques conclusives	300
5.2	<i>Mécanique statistique</i>	302
5.2.1	Formalisme	303
5.2.2	Entropie de Boltzmann	307
5.2.2.1	Entropie continue de Boltzmann	307
5.2.2.2	Entropie combinatoire de Boltzmann	311
5.2.2.3	Entropie macroscopique de Boltzmann	315
5.2.3	Entropie de Gibbs	317
5.3	<i>Discussion</i>	323
5.3.1	Connexions	323
5.3.2	Réduction	324
5.3.3	Irréversibilité	325
5.3.4	Interprétations	330
5.3.4.1	Probabilité	331
5.3.4.2	Désordre	337
5.3.4.3	Ignorance et information	341
5.3.4.3.1	Machine de Szilard et principe de Landauer	345
5.3.4.3.2	Théorie de l'information	349
5.3.4.4	Travail	351
5.3.4.5	Dispersion	355
5.3.5	Remarques conclusives	356
5.3.6	Stratégie explicative	359
5.3.6.1	Équilibre	360

5.3.6.2	Augmentation d'entropie	363
5.4	<i>Autres disciplines</i>	371
5.4.1	Théorie de l'information	371
5.4.2	Mécanique quantique	372
5.4.3	Biologie	375
5.5	<i>Conclusion</i>	378
6	ENTROPIE : EXPLANANDUM OU EXPLANANS DE L'ÉMERGENCE ?	381
6.1	<i>Systèmes complexes</i>	384
6.2	<i>L'entropie comme explanandum</i>	388
6.2.1	Quelle entropie ?	389
6.2.2	Quelle émergence ?	393
6.2.3	Quel explanandum ?.....	396
6.2.4	Discussion & conclusion.....	399
6.3	<i>L'entropie comme explanans</i>	403
6.3.1	Approche substantielle de l'entropie comme explanans.....	405
6.3.1.1	Stratégie généralisée et systèmes complexes	406
6.3.1.2	Probabilités, interventions et entropie.....	412
6.3.1.3	Physiologie	415
6.3.1.4	Théorie constructale	417
6.3.1.5	Remarques conclusives	420
6.3.2	Approche analogique de l'entropie comme explanans	421
6.3.2.1	Probabilités, information et entropie.....	423
6.3.2.2	Génétique des populations	425
6.3.2.3	Éconophysique	430
6.3.2.4	Remarques conclusives	432
6.3.3	Discussion & conclusion.....	432
6.4	<i>Conclusion</i>	437
7	CONCLUSIONS GÉNÉRALES	443
7.1	<i>Le paradoxe du chapeau et de la fourmi</i>	443
7.2	<i>Complexité et probabilités</i>	445
7.3	<i>Un pluralisme restreint</i>	450
7.4	<i>Une polysémie ambitieuse et féconde</i>	455
7.5	<i>L'émergence méthodologique</i>	465
7.6	<i>L'entropie comme explanandum et explanans de l'émergence</i>	468
7.7	<i>Multi-, inter- et transdisciplinarité</i>	471
8	BIBLIOGRAPHIE	473
9	PRÉSENTATION DE L'AUTEUR	492

À Gabrielle,
Grégoire et Laure

REMERCIEMENTS

Je tiens à remercier mes co-directeurs, Anouk Barberousse et Frédéric Bouchard, qui ont su former une équipe interdisciplinaire hors du commun. Je vais continuer à travailler les commentaires toujours pertinents d'Anouk sur les versions préliminaires de cette thèse, et je vais retenir les conseils stratégiques redoutables formulés par Frédéric.

Je veux aussi remercier Philippe Huneman pour le séminaire que j'ai suivi à Paris à l'hiver 2011, ainsi que Thierry Martin qui m'a accueilli au sein de son laboratoire lors d'un stage à Besançon à l'hiver 2013. Je remercie aussi Mario Bunge et Jean-Pierre Marquis pour leur aide et leurs conseils.

Mes remerciements vont aussi au Fonds de recherche du Québec sur la société et la culture pour avoir (enfin) compris l'intérêt de cette recherche et de m'avoir soutenu financièrement pendant un peu plus d'un an.

Je reconnais aussi le soutien financier de l'université de Montréal et du Centre interuniversitaire de recherche sur la science et la technologie.

Merci à Catherine & Oscar, Renée & Vincent, et Fabienne & Pierre pour m'avoir hébergé lors de mes séjours à Montréal. Merci aussi à Annie & Bruno et Véronique & Étienne pour leur soutien auprès de ma famille lors de mes déplacements à Paris.

Je remercie aussi mes parents, qui m'ont soutenu dans ce projet. Chacun, à sa façon, a su apporter son aide au moment opportun.

Enfin, je tiens à remercier tout spécialement ma compagne de vie, Gabrielle, qui sait partager mes joies comme mes peines, qui sait être forte dans l'épreuve, et qui a su montrer une incroyable patience au cours de ce projet, je dois le dire, un peu fou. Mes pensées vont aussi bien sûr à mes enfants, Grégoire et Laure, qui m'ont appuyé sans le savoir. Puissent ces efforts ne pas être vains et leur enseigner que la rançon de la gloire se juge à l'aune de ces efforts.

1 Introduction générale

RÉSUMÉ. Le concept d'entropie fait paradoxalement consensus sur deux points : il est fondamental et personne ne sait vraiment de quoi il s'agit. Sa polysémie ne semble pas un obstacle puisqu'il est mobilisé au sein d'explication de phénomènes courants comme la tendance à l'équilibre thermique, et plus complexes comme l'émergence de structures dissipatives, voire l'émergence de la vie, la diffusion de l'information, et même la direction du temps. Or un examen de ces exemples variés montre des usages très différents tant de ce concept que des notions mêmes d'explication et d'émergence. Ainsi, le concept d'entropie se présenterait comme ce que permet d'expliquer le concept d'émergence (comme inexplicabilité du tout par les parties), mais aussi, comme l'élément central d'une explication de l'émergence de nouvelles structures dans les systèmes complexes. Si l'explication est une notion asymétrique, alors l'entropie ne peut être à la fois explanans et explanandum de l'émergence. Encore faut-il que le sens du concept d'émergence soit clair. Il y a là matière pour le philosophe.

Les chemins qui vont droit devant soi s'escaladent.

Gilles Vigneault

Un professeur qui demandait sempiternellement la même question à l'examen final, à savoir : « qu'est-ce que l'entropie ? », se vit suggérer par son assistant de modifier la question afin de placer les étudiants dans l'expectative ou à tout le moins de créer un peu de nouveauté. À cela le professeur répondit : « c'est toujours la même question, mais la réponse change à chaque fois ».

Et le sage d'ajouter : il n'y a pas de mauvaises réponses, que de mauvaises questions. Ici la question demande cependant un préambule commençant au milieu du XIX^e siècle dans une chaudière à charbon, pour ensuite passer par des équations différentielles, s'arrêter dans un laboratoire de télécommunication, avant de poursuivre sa route parsemée d'arrêts impromptus parce que quelqu'un, quelque part, croit reconnaître des ressemblances avec quelque chose. On l'aura compris : le concept d'entropie est polysémique.

Il est pourtant fondamental. Dans sa version originale issue de la thermodynamique, il est défini, plutôt prosaïquement, par le ratio de la quantité de chaleur échangée par la température de cet échange. Le principe entropique stipule que cette quantité ne peut qu'augmenter dans le temps et il exprime l'impossibilité du mouvement perpétuel de second type, où une quantité de chaleur serait entièrement convertie en travail. Comme la chaleur est pour ainsi dire partout, il met donc un terme à

ce désir sans doute pathétique d'obtenir quelque chose à partir de rien. En même temps, il semble rendre compte de certaines intuitions tenaces, comme ces croyances en l'impossibilité de reformer la fumée et la cendre en arbre ou de désembrouiller les œufs d'une omelette.

Deux physiciens de renom, Planck (1945) et Eddington (1928), plaçaient ce principe en très haute estime, le premier le considérant comme rien de moins qu'une « vérité apriorique » et le second le voyant à l'abri d'une réfutation théorique. Deux philosophes aussi reconnus, Reichenbach (1956) et Carnap (1977), lui ont réservé des études approfondies et attribué des vertus explicatives considérables.

L'une de ces vertus se manifeste dans l'explication des systèmes dits complexes. Ces systèmes ont la particularité de rendre la méthodologie classique, consistant à expliquer de grands systèmes par l'analyse de ses parties et de leur association selon un schème cumulatif, particulièrement difficile. Il semble évident, par exemple, que le métabolisme d'un mammifère ne puisse être expliqué par les lois de la mécanique quantique. Mais ces difficultés ne se limitent pas au phénomène de la vie. Certains systèmes physiques, dits loin de l'équilibre, échappent à cette possibilité d'une décomposition simple du tout par les parties, de réduire le tout à la somme des parties. Ces systèmes adoptent dans certaines conditions une organisation structurale qui, justement, peut être fort complexe. On dit alors que l'unité de cette organisation, cet « ordre », émerge de la dynamique des sous-unités.

Or, les conditions qui font en sorte que ce genre de comportement soit possible concernent, justement, l'entropie. C'est-à-dire que lorsqu'un système absorbe de l'énergie et qu'il dissipe dans son environnement une certaine quantité d'entropie, alors s'offre à lui la possibilité de s'« auto-organiser ». De sorte que l'explication de ce type de comportement, soit l'émergence d'une telle organisation, mobilise à titre d'élément primordial le concept d'entropie. En termes simples, l'entropie ou son comportement serait la « cause » de ce genre de phénomène émergent, l'entropie expliquerait l'émergence. Les précisions et les nuances que cette brève présentation exige constituent un élément majeur de cette thèse.

Mais ce n'est pas le seul lien que le concept d'entropie entretient avec ceux d'émergence et d'explication. Car nonobstant ces vertus explicatives, la polysémie du concept d'entropie fait toujours rage. La raison pour laquelle on retrouve autant de définitions de l'entropie trouve son origine dans la tentative d'une traduction de sa version thermodynamique dans le langage de la mécanique classique, avec le concours des probabilités et des statistiques, d'où la mécanique statistique. Le résultat est un concept abstrait, dont la réelle signification fait l'objet de multiples interprétations, et qui semble pouvoir s'appliquer à plusieurs contextes. Mais les résultats ne sont pas aussi convaincants qu'on le voudrait bien et il n'est pas clair que la traduction soit bien effectuée dans le langage ciblé au départ, celui de la mécanique. Autrement dit, la réduction d'un concept à l'autre ne semble pas opérer.

Comme précédemment, il y aurait, croit-on, émergence des propriétés thermodynamiques comme l'entropie par rapport aux propriétés des constituants des systèmes typiquement thermodynamiques, soit les molécules. Plus précisément, la mécanique statistique serait incapable de reproduire l'irréversibilité thermodynamique, et par conséquent, l'entropie serait émergente, ouvrant la voie à l'explication de l'irréversibilité thermodynamique par l'émergence.

Alors, est-ce possible que l'entropie soit *à la fois* une propriété émergente *et* ce qui cause l'émergence ? Le concept d'entropie peut-il être à la fois explanandum et explanans (ou plutôt : son élément principal) de l'émergence ?

Je vais montrer dans cette thèse sous quelles conditions cela est possible et fournir ainsi un cadre explicatif de phénomènes divers et complexes, de la physique à la biologie. Je soutiens que l'analyse est plus fructueuse si ces deux questions sont traitées de front. Mais pour y répondre, il importe de tenir compte de plusieurs subtilités et d'éviter certains pièges, ce qui exige, dans un effort pédagogique soutenu, une discussion rigoureuse et approfondie sur plusieurs enjeux de la philosophie des sciences.

1.1 Philosophie des sciences

1.1.1 Mise en contexte

La *Rede Lecture* de 1959 à l'université d'Oxford a d'abord été l'objet d'un article dans un numéro de 1956 du *New Statesman*, par Charles Percy Snow. Un mois après la publication de cet article, le nouvelliste et professeur de biochimie américain Isaac Asimov terminait *The Last Question*, une courte nouvelle dont le thème central était l'impérieuse réalité de l'augmentation de l'entropie à son maximum – la « fin de la partie » en quelque sorte. Elle raconte l'histoire de l'humanité poursuivant son osmose avec la technologie, alors que la Terre était alimentée par l'énergie solaire et que les ordinateurs s'ajustaient et se corrigeaient eux-mêmes, chaque génération posait inlassablement la même question : « Comment l'entropie de l'univers peut-elle être massivement réduite ? », et de recevoir invariablement comme réponse : « Il n'y a pas suffisamment de données pour une réponse signifiante ». Une fois l'humanité éteinte, une fois la somme de son potentiel mental était encodée dans un superordinateur, AC, les étoiles s'effondraient les unes après les autres et les planètes se refroidissaient. L'énergie et la matière s'étaient évanouies, et avec elles le temps et l'espace. AC n'existait que pour cette dernière question qu'il avait enfin réussi à résoudre. Mais il n'y avait personne à qui transmettre cette réponse : l'humanité et l'univers s'étaient éteints. La conscience de AC englobait tout ce qui avait jadis été un univers et il repassa sur ce qui était désormais le chaos. Il pensa à comment transmettre sa réponse : « Que la lumière soit », dit AC. Et la lumière fut.

La morale eschatologique de cette nouvelle ne paraît pas accessible à tous. Mêler ainsi le destin de l'humanité à un principe thermodynamique renvoie dos à dos science et littérature. Or c'est précisément de cette prévisible inaccessibilité de la morale, et non la morale elle-même, dont se lamentait Snow dans son article. Il déplorait ainsi l'« incompréhension mutuelle » séparant la culture scientifique de la culture dite traditionnelle. Il décrivait cette dernière comme essentiellement littéraire, se comportant tel un État dont le pouvoir décline rapidement, et « dépassant beaucoup trop d'énergie sur des subtilités alexandrines ». Ne pas avoir lu *Guerre et paix* ou la *Chartreuse de Parme*¹ revient à ne pas être éduqué, mais il en est tout autant de ne pas avoir la moindre idée du second principe de la thermodynamique. Ce serait l'équivalent scientifique de « Avez-vous lu un ouvrage de Shakespeare ? » Le bris de communication entre ces deux cultures, soutenait-il, constituait un obstacle majeur à la résolution des problèmes du monde.

¹ Ironiquement, Stendhal a dédié son roman « *To the Happy Few* » (en anglais), sorte de cénacle littéraire.

La philosophie des sciences a cette ambition tacite d'instaurer cette communication, d'établir des ponts entre les cultures traditionnelle et scientifique. Qu'elle soit en mesure de résoudre les problèmes du monde, en revanche, relève davantage de la prétention des philosophes que de leur réelle envergure. Pour qu'il y ait véritablement dialogue, bien entendu, l'échange ne doit pas tourner au monologue ni au soliloque. Une philosophie qui se coupe des sciences n'avance pas, et une science sans philosophie est vide. Or « science sans conscience n'est que ruine de l'âme » nous rappelle Rabelais. Mais si la philosophie se contente d'écouter ce qu'ont à dire les sciences, elle repartira avec des problèmes qui ne sont pas les siens. C'est pourquoi elle ne peut faire l'économie d'un discours *normatif* sur les sciences. Ainsi, un discours prétendant offrir un critère de démarcation entre science et non-science, ou même une taxonomie des sciences, verse déjà en philosophie. Et cela vaut aussi pour le scientifique condescendant envers le philosophe. La philosophie a en effet cette particularité d'être automatiquement pratiquée par ceux qui en parlent : ceux et celles croyant pouvoir s'en passer en toutes choses en font généralement de la mauvaise.

La particularité de la philosophie des sciences, plus spécifiquement, est d'être par nature interdisciplinaire, au sens où elle coordonne des concepts généraux inspirés et, en retour, appliqués à plusieurs disciplines. Bien que le terme « interdisciplinaire » soit un enfant du siècle dernier, il résonne au cœur d'idées comme la science unifiée, la connaissance générale, la synthèse et l'intégration des connaissances. Alors que de nouveaux champs de connaissances apparaissent continuellement, avec leur objet propre, les méthodes, les instruments et les concepts sont transposés d'un champ à l'autre : même ceux qui prônent la désunion des sciences ont du mal à expliquer cette interdisciplinarité. Selon Platon, le philosophe est le plus à même de synthétiser, de systématiser les connaissances. Et son élève Aristote, malgré quelques dissonances avec son illustre maître, accordait aussi au philosophe cette capacité d'embrasser toutes les disciplines du savoir, au sens générique et étendu du terme. Or, depuis la réforme universitaire prussienne ayant tourné le diplôme de doctorat vers des recherches nouvelles dans une discipline reconnue, la surspécialisation est devenue la norme : « *to learn more and more about less and less* ». La philosophie des sciences ne fait pas exception. Le phénomène est encore plus marqué au sein des disciplines connexes, comme la philosophie de la physique ou de la biologie. Les récentes contributions dans ces domaines comme dans d'autres se résument bien souvent à poser une brique sur un édifice dont on a perdu la vue d'ensemble. L'intérêt récent pour l'interdisciplinarité est ainsi un intérêt renouvelé qui se distingue de ses motivations originales sans toutefois les rejeter complètement. Nul doute que les appels de Platon et d'Aristote résonnent encore.

Le dialogue peut et je dirais doit être fructueux. Si la connaissance est le pouvoir, comme le dit Bacon, encore faut-il l'exercer avec discernement. Surtout, ajouterai-je, lorsque l'une ou l'autre fait défaut. Comme le suggère Asimov, peu importe l'avenir, il y a de la lumière à la fin.

1.1.2 Complexité

Le discours normatif que la philosophie peut porter sur les sciences est essentiellement de trois ordres : métaphysique, épistémologique ou éthique. Ce faisant elle ne doit pas faire la sourde oreille au discours scientifique. Il n'est plus question maintenant de se couper de l'expérience afin d'établir l'inventaire du monde, de dire en quoi constitue la réalité. Et il est devenu commun de motiver une théorie de l'explication ou de la connaissance en se basant sur la méthodologie ou les concepts de certaines disciplines scientifiques (et même de s'en inspirer pour l'attribution de valeurs morales). Or, des développements récents en sciences viennent remettre en question plusieurs conceptions métaphysiques et épistémologiques, comme celles d'individuation, de prédictibilité, de déterminisme, de causalité, de réduction, d'émergence et de lois de la nature. Ces développements concernent principalement ce qui est désormais convenu d'appeler la complexité et les systèmes complexes.

De toute évidence, déterminer en quoi consiste la réalité, et pour peu que l'on s'entende sur ce que cela signifie, n'est pas une mince affaire. Ce n'est pas pour rien que certains scientifiques posent un regard empli de modestie sur leur discipline et que plusieurs philosophes ont été amenés à nier la possibilité même de le faire. Cette possibilité renvoie nécessairement au procès de la connaissance : qu'est-ce que savoir, quoi et comment ? Comme ces questions sont souvent entremêlées, elles prennent un aspect particulièrement délicat dans le cas des systèmes complexes présentant divers niveaux d'organisation, des atomes aux galaxies en passant par les cellules vivantes et les écosystèmes. Car il est patent que les systèmes à divers niveaux ne sont pas tous également accessibles, certains étant observables et d'autres non, et qu'ils présentent divers degrés de prédictibilité et de complexité, au sens du nombre d'interrelations entre sous-systèmes – le « *complexus* », pour Edgar Morin, est ce qui est « tissé ensemble ». Par exemple, je puis assez aisément prédire le refroidissement de ma tasse de café, mais je serais bien en peine de prédire la trajectoire ne serait-ce que d'une seule molécule du contenu de la tasse. Or, un discours sur cette molécule, ou même sur l'ensemble des molécules de la tasse, est-il nécessairement illégitime, non fondé épistémologiquement ? Tout dépend du discours, bien entendu. Dire que la molécule existe est une chose, décrire son comportement en est une autre, de même que, et surtout, affirmer que son comportement *doit* être tel, autrement dit de l'expliquer. Démêler ces assertions quant à leur nature, séparer le bon grain de l'ivraie, exige un travail subtil. Face à ces difficultés, deux impératifs semblent se dessiner : d'une part, éviter un réalisme naïf stipulant la vérité absolue de nos croyances sur les apparences ou la correspondance terme à terme entre la réalité et la théorie, et d'autre part, adopter la méthodologie visant à distinguer à des fins d'analyse ces divers niveaux d'organisation (Wimsatt 1994, 2000 et 2006 ; Lenk & Stephan 2002 ; Bunge 2006).

Ces impératifs devront toutefois recevoir des justifications. D'où l'importance de situer la discussion dans un large contexte épistémologique. Toute discussion portant sur les niveaux implique un ensemble de problèmes et engage des débats importants et persistants en philosophie des sciences. On peut par exemple avancer que les propriétés *en apparence* différentes d'un système composé (tout, unité) ne sont pas *en réalité* différentes de celles de ses composants (parties, sous-unités). Que ses propriétés soient en définitive, tout bien considéré, différentes de celles des composants est généralement inféré de l'impossibilité d'en donner une explication en termes des lois régissant les propriétés des composants. La notion d'explication, pour autant qu'elle soit claire, joue alors le rôle d'un paramètre évaluant la pertinence et le statut de diverses relations inter-théoriques. S'il était possible, par exemple, d'expliquer la biologie par les lois de la physique, alors la relation que l'une entretiendrait avec l'autre serait celle d'une *réduction*, et la biologie ne serait rien d'autre que de la physique déguisée, en quelque sorte. À cet enseigne, toutefois, on est très loin du compte et la biologie a de très beaux jours devant elle. Pour autant, on présume que les systèmes biologiques sont tout de même constitués d'entités physiques, comme les atomes, dont la description est bien sûr du ressort de la physique. Cette situation exprime aussi une relation inter-théorique, celle de l'émergence. Le concept d'émergence est ainsi censé offrir un compromis entre une position franchement réductionniste, ou tout ne serait que physique ou « idée », et une position pluraliste séparant radicalement différents niveaux, comme la matière et l'esprit (ce concept est introduit plus loin).

Ces considérations épistémologiques définissent le cadre d'étude des systèmes complexes (Chapitre 6), c'est-à-dire des systèmes constitués d'un grand nombre d'éléments interagissant de manière non triviale ou non simple, selon la conception classique de Simon (1962). En effet, ces systèmes, point focal de plusieurs champs de recherches souvent interdisciplinaires, peuvent être étudiés à plusieurs niveaux, de complexité variable et faisant l'objet de diverses théories scientifiques, de la géographie physique à la socio-économie. Ces systèmes sont généralement associés à des notions telles que la non-linéarité, la non-additivité, l'irréversibilité, la brisure de symétrie, l'ordre, l'émergence, la hiérarchie, l'historicité, etc. (Hooker 2011). Les systèmes complexes présentent ainsi, généralement, des composantes qui sont, à un certain degré ou à un autre, indépendantes, et donc autonomes dans leur comportement, tout en subissant diverses interactions ; il en résulte un comportement global du système difficilement prévisible ou explicable sur la base de la connaissance de ses composants. L'analyse de tels systèmes se penche sur ce qui est donc appelé des *niveaux d'organisation* (Chapitre 2) présentant des caractéristiques suffisamment stables pour être étudiées, mais qui peuvent néanmoins être « entremêlés » et donc présenter des obstacles à cette analyse. C'est-à-dire que la méthode classique, consistant à ramener le complexe au simple selon des modèles linéaires et déterministes, montre des signes de faiblesses, même si elle est parfois forcée, malgré tout.

Ces caractéristiques des systèmes complexes constituent ainsi une nouveauté à laquelle font face diverses disciplines scientifiques et qui a su déstabiliser l'approche réductionniste longtemps dominante. C'est pourquoi la notion d'émergence fait à l'heure actuelle l'objet de plusieurs discussions philosophiques. Cette nouveauté, voire ce nouveau paradigme pour certains (par ex. Prigogine & Stengers 1979), est tributaire d'une vaste révolution épistémologique débutée au XIX^e siècle avec la thermodynamique classique et qui s'est poursuivie durant la seconde moitié du XX^e siècle avec la thermodynamique du non-équilibre (Chapitre 3). Mais il est loin d'y avoir eu stagnation entre les deux. Dans les années 1960, Ed Lorenz, travaillant à un modèle de prévision météorologiste, procéda à une simulation qui donna un résultat fort différent d'une simulation précédente avec pourtant le même modèle. La différence résidait en ceci que Lorenz avait démarrée la seconde simulation en choisissant comme conditions initiales un point (l'état du système) particulier de la première simulation, mais en omettant malgré lui d'inscrire les dernières décimales. Cet écart entre les résultats des simulations a été représenté comme étant l'équivalent d'un écart entre un ouragan et une journée de beau temps, et cette infime différence dans les conditions initiales, métaphoriquement, comme le battement d'aile d'un papillon – d'où l'« effet papillon » désignant la sensibilité aux conditions initiales. Un peu plus tard dans les années 1970, David Ruelle et Floris Taken ont affublé du vocable « chaos » une certaine classe de solutions d'équations différentielles partielles particulièrement difficiles, tandis que Benoît Mandelbrot présentait sa « géométrie fractale » et René Thom sa « théorie des catastrophes ». On peut parler d'une période de grands changements – peut-être d'un « changement de paradigme scientifique structural » (Schmidt 2011) – où l'instabilité n'était plus vue comme un sous-groupe de la stabilité, mais plutôt l'inverse.

Le concept d'entropie est issu de cette vaste révolution épistémologique, à son début plus spécifiquement. Ce concept, émanant d'abord de la thermodynamique (Chapitre 4) puis reformulé en mécanique statistique (Chapitre 5), apparaît particulièrement intéressant pour l'étude de la relation entre niveaux d'organisation de systèmes complexes (Chapitre 6). En effet, la mécanique statistique est une théorie tentant d'expliquer le comportement observable (niveau macroscopique) de grands systèmes à la dynamique de molécules inobservables (niveau microscopique). En quelque sorte, donc, cette discipline a toujours traité d'émergence. L'entreprise visant à faire sens du concept d'entropie et des prétentions explicatives le mobilisant est donc aux prises avec un réseau de problèmes épistémologiques de taille, allant aux fondements même de la connaissance (évoqués plus tôt).

1.2 Les 3 *E*, la fourmi et le chapeau

En guise d'introduction à la présentation du projet global, aux prises avec ces trois concepts qui ne se laissent pas aisément appréhendés – les « 3 *E* » : l'entropie, l'émergence et l'explication – je propose une petite fable, question de rendre le tout un peu plus digeste, ou du moins de dresser un portrait global mais succinct de la situation.

D'un physique plutôt ingrat, une fourmi était convaincue qu'un chapeau lui siérait bien. Après avoir muri sa réflexion quant à la forme et à la couleur idéales, son choix s'arrêta sur un chapeau melon. Malheureusement, ou heureusement selon le point de vue, le chapeau faisait trois fois sa taille : elle disparue sous celui-ci, mais persuadée qu'il lui faisait à ravir. Tant et aussi longtemps qu'elle portait ou semblait porter son chapeau, nul ne la voyait. Pour se pavaner un peu à la vue des bonnes gens, il lui fallait quitter son nouveau nid. Il lui fallait retirer son nouveau chapeau.

Métaphoriquement, la fourmi est l'entropie thermodynamique et le chapeau l'entreprise réductionniste de l'entropie statistique : en désirant reformuler le concept d'entropie thermodynamique, de le vêtir sous de nouveaux appareils, le concept original a disparu. Du moins, il a drôlement changé. Que cela soit ou non une bonne chose dépend en grande partie des desiderata d'une théorie scientifique car plusieurs dédaignent une science phénoménologique, des principes qui ne sont pas mathématisés, des quantités physiques qui demandent la définition de l'environnement du système ou une stratégie explicative se basant sur un cas limite jamais instancié. Certains pensent ainsi que la fourmi a une physionomie qu'il est préférable de cacher (sauf qu'ici la fourmi ne représente pas le microscopique mais le macroscopique). Le concept a changé car il provient bien sûr d'un autre contexte théorique et ce faisant il a perdu son caractère d'irréversibilité. Avec son chapeau, la fourmi ne peut se déplacer : dans sa version statistique, l'entropie n'augmente pas. Pour explorer le monde, la fourmi doit quitter son chapeau : pour expliquer de nouvelles structures complexes, l'entropie doit laisser de côté sa version statistique. Pourtant, comme on le verra, le chapeau semble faire à plusieurs, car le concept d'entropie statistique a été récupéré dans plusieurs disciplines, comme la génétique et l'économie.

Il paraît que le mérite de la fable est double : elle suscite le rire et donne une leçon de prudence. Les trois concepts formant les 3 *E* sont couramment mobilisés au sein de discussions sur le thème générique de la complexité. Il n'y a peut-être pas de quoi rire, mais nous sommes prévenus.

1.2.1 Entropie

On a vu que, paradoxalement, le concept d'entropie fait consensus sur deux points : c'est un concept fondamental et personne ne sait vraiment de quoi il s'agit. Comme le soulignait Khinchin déjà en 1949 (137), très peu de notions scientifiques ont reçu autant d'efforts afin de clarifier leur signification théorique et philosophique. Il affiche de grandes ambitions dont la réalisation laisse perplexe, alors qu'il révèle une fécondité philosophique et scientifique rarement égalée.

Alors comment résoudre cette ambivalence, voire cet apparent paradoxe ? Ici aussi la question demanderait un long préambule, mais il est possible de simplifier sans dénaturer le problème. La polysémie du concept d'entropie origine d'abord dans ce qu'on pourrait appeler, en première approche, une tentative de traduction. L'avènement d'une nouvelle discipline rendant compte des transformations de chaleur et de travail, de même qu'une loi stipulant l'augmentation d'une quantité physique dans le temps, en porte-à-faux des autres lois physiques, ont poussé certains physiciens à tenter de ramener l'inconnu au (plus) connu, c'est-à-dire de traduire, si faire se peut, la thermodynamique à la mécanique. Devant le nombre gigantesque d'atomes ou de molécules constituant le moindre corps susceptible d'être observé et d'intéresser la thermodynamique, la mécanique classique se trouvait dans l'embarras et devait faire appel aux statistiques et aux probabilités. L'entreprise n'est ni vaine ni illégitime, mais elle a produit un nouveau concept d'entropie, plus abstrait, et irrémédiablement lié à la notion de probabilité qui, au demeurant, ne se laisse pas aisément appréhender pour peu qu'on l'y regarde de plus près. De là origine sa polysémie, et il a par la suite été récupéré par moult disciplines, de la biologie à l'économie en passant par les télécommunications. (L'historique du concept est l'objet du Chapitre 3.)

Cette polysémie, voire cette ambiguïté, n'enlève rien à ses prétentions. L'entropie est en effet le centre d'explications de phénomènes divers comme cette tendance notoire des tasses de cafés à se refroidir d'elles-mêmes et la nécessité de payer pour l'énergie qui pourtant se conserve, et vient même appuyer des récits eschatologiques comme dans la nouvelle d'Asimov. Ce caractère singulier de son augmentation dans le temps lui vaut une place privilégiée au sein des discussions sur la nature du temps et plus particulièrement sur sa directionnalité. L'objectif premier de ces discussions est d'expliquer pourquoi certains processus semblent se produire dans une seule direction, pourquoi ils présentent une asymétrie temporelle, comme le vieillissement, par exemple. La raison pourrait être la nature du temps lui-même ou plutôt un principe ou une loi régissant ce genre de processus qui sont dits irréversibles. Deux sens de la notion d'irréversibilité doivent alors être distingués, la non-invariance par renversement (inversion de la séquence des états) et l'irrécupérabilité (impossibilité de retrouver l'état

initial). Le second principe de la thermodynamique et le principe entropique, on le verra, exemplifient ces deux sens.

L'origine du concept d'entropie remonte à la révolution industrielle alors que les ingénieurs tentaient de comprendre et d'optimiser le rendement des machines à vapeur. Les travaux de Sadi Carnot, Polytechnicien, s'appuyaient sur l'impossibilité du mouvement perpétuel et ont constitué un élément majeur dans l'heuristique du concept d'entropie en ce qu'ils présentent le concept de cycle réversible duquel une quantité maximale de travail peut être extraite d'une quantité de chaleur donnée. Mais ils montrent surtout que la production d'une quantité de travail à partir d'une quantité de chaleur est limitée de principe, ce qu'exprime en substance le second principe de la thermodynamique, que je distingue du principe entropique, selon lequel l'entropie d'un système isolé ne peut qu'augmenter. J'offre une reconstruction conceptuelle du second principe à partir de ces concepts de chaleur et de travail, mais aussi de différence de température (Chapitre 4).

La formalisation mathématique ainsi que le nom de l'entropie (du grec « transformation » et du préfixe « en ») sont dues au physicien allemand Rudolf Clausius. Elle a comme ambition de reformuler ce principe pourtant fondamental qu'est l'impossibilité du mouvement perpétuel. L'idée générale derrière cette formalisation est qu'une quantité $-\delta Q/T$ se conserve lors d'un processus réversible entre deux états d'équilibre mais ne se conserve pas lors d'un processus irréversible. Comme la tendance à l'équilibre des systèmes isolés de leur environnement constituait déjà un fait physique relativement bien établi (Fourier), cette quantité est censée augmenter puis se maintenir dans le temps pour tout système isolé de son environnement et atteignant l'équilibre après un certain temps. Or, un processus réversible est une idéalisation qui n'est donc jamais instanciée. Tous les processus ne sont donc pas réversibles, et sont donc irréversibles. Le concept d'entropie est ainsi articulé au sein d'un principe, le principe entropique, stipulant l'augmentation de l'entropie dans le temps. Il importe alors de noter que, malgré la possibilité d'une diminution locale d'entropie, l'entropie globale doit augmenter, en sorte qu'une diminution est toujours surcompensée par une augmentation dans l'environnement du système. L'entropie d'un système isolé, sans échange d'énergie ni de matière, ne peut donc qu'augmenter jusqu'à ce qu'il atteigne l'équilibre. Sortir le système de cet état d'équilibre nécessiterait une intervention de l'extérieur du système. Ce principe est donc censé fournir une explication des processus vers l'équilibre et donc irréversibles *parce que* l'entropie *doit* augmenter.

Plusieurs critiques ont été adressées au second principe de la thermodynamique ainsi qu'au principe entropique. Dans le premier cas on reproche un principe « négatif » ne statuant que sur ce qui impossible, qu'il s'appuie sur un cas limite jamais instancié, soit un processus réversible, ou encore qu'il ne décrit pas l'état des systèmes et doit faire appel à la description de son environnement. Dans le

second cas, les reproches sont similaires, sans s'y réduire, car on lui reproche ses idéalizations, des processus réversibles et à l'équilibre, mais aussi de s'appuyer sur une fonction d'état qui n'est définie que pour un cas limite jamais instancié, et qu'il ne peut décrire ni expliquer les phénomènes irréversibles. Mais on verra que ces objections n'affectent pas la prétention à la validité scientifique de ces principes, pour peu qu'ils soient semblables à d'autres principes acceptés, ni leur vertu explicative. La critique qui pourrait être fatale cependant est celle interdisant, pour quelques raisons épistémologiques, l'emploi d'un cas limite jamais instancié au sein de l'explication.

L'ambition de la mécanique statistique n'est pas moindre, car elle entreprend de reformuler dans un autre contexte théorique le concept d'entropie et le principe entropique, ce dernier étant déjà la reformulation d'un principe jugée fondamentale, voire inébranlable. Les travaux de pionniers comme Clausius, Maxwell et Boltzmann tentaient d'expliquer les systèmes typiquement thermodynamiques par des procédures statistiques appliquées à de grands ensembles de particules. Mais les paramètres dits macroscopiques, puisqu'ils sont observables ou mesurables, ne spécifient pas l'état dit microscopique déterminé par l'ensemble des paramètres dynamiques pour chacune des particules. Ainsi, plusieurs états microscopiques sont compatibles avec un seul état macroscopique. La stratégie orthodoxe de formalisation en mécanique statistique consiste ainsi à définir une fonction d'une dynamique sur l'espace des phases canonique. Autrement dit, elle repose sur l'idée que le comportement des systèmes macroscopiques ou phénoménologiques, bref thermodynamiques, dépend d'une manière ou d'une autre de la dynamique microscopique de leurs constituants que sont les particules, les molécules. Cependant, cette dynamique n'est pas tractable pour les systèmes d'une grandeur suffisante pour être justement considérés comme macroscopiques et c'est pourquoi l'approche statistique est adoptée. C'est-à-dire qu'il ne s'agit pas de retracer chaque molécule afin d'en décrire l'histoire mais plutôt d'en déterminer le comportement global en calculant des valeurs statistiques, comme des moyennes. Cette méthode requiert des hypothèses simplificatrices qui posent des problèmes conceptuels d'envergure quant à la validité des résultats de cette approche.

Or, l'entropie statistique est en quelque sorte une quantification de cette compatibilité. Il n'est donc pas surprenant, compte tenu de ce rapport « un-plusieurs » entre un état macroscopique et des états microscopiques, que l'entropie ait été associée à la probabilité de cet état macroscopique. Certains ont même avancé, dont Boltzmann lui-même, compte tenu de l'augmentation d'entropie stipulée par le principe entropique, que les systèmes thermodynamiques passaient d'états peu probables à des états plus probables. C'est pourquoi ce chapitre s'ouvre sur une importante discussion sur les probabilités. Non pas sur le calcul des probabilités, ou sur leur axiomatisation, mais plutôt sur leur signification et leur interprétation en physique. Je présente d'ailleurs six définitions différentes de l'entropie, quatre en mécanique statistique selon les approches de Boltzmann et Gibbs, et celles de la mécanique quantique

et de la théorie de l'information, suivies de cinq interprétations différentes (Chapitre 5). C'est pourquoi on peut bien soutenir qu'il n'y a « pas de définition complètement satisfaisante de l'entropie » (Càpek & Sheehan 2005 : 23).

Le concept d'entropie est donc rempli de promesses mais reste à mains égard ambigu. D'où l'intérêt d'un examen critique. Mes efforts d'analyse se sont concentrés sur les concepts de la thermodynamique et de la mécanique statistique. Je n'ai pas traité, par exemple, de l'entropie non-extensive (Tsallis 1988, 2001 et 2002 ; Tsallis & Brigatti 2004), qui se veut une généralisation de l'entropie statistique aux systèmes dont la dynamique est plus complexe, ni de certaines subtilités de la mécanique quantique qui demanderaient une étude encore plus étendue. J'ai tout de même confiance que les deux chapitres consacrés à cet examen critique de ce concept soient suffisamment complets. On verra que l'entropie associée à des concepts d'intérêts philosophiques, comme la direction du temps et l'irréversibilité, la causalité, la probabilité et la structure des théories et des explications scientifiques. Il intervient aussi au sein d'explications scientifiques particulières comme celles décrivant un processus vers un état d'équilibre, ou celles définissant la probabilité d'un état microscopique à partir d'un état macroscopique. Dès lors qu'un concept, ou son « étiquette », reçoit plusieurs significations et interprétations, la question est de savoir si l'une ou plusieurs d'entre elles peuvent en expliquer d'autres.

1.2.2 Explication

Au sens le plus large, une explication, qu'elle soit scientifique ou non, est censée réaliser plusieurs tâches : montrer que ce qui s'est produit devait se produire ou qu'on était en droit de s'y attendre, faire comprendre, unifier plusieurs phénomènes semblables, permettre de prédire, de contrôler, de manipuler. Bien que cette liste ne soit pas exhaustive, on constate qu'il y a plusieurs recoupements entre ce que l'on cherche d'une explication et ce que cherche à obtenir le scientifique en général, ce qu'il poursuit ou juge avoir une certaine valeur dans son travail. Avec une portée aussi large, il paraît naturel de dire que le « but de la science » consiste à générer des « explications satisfaisantes » (Popper 1991 : 297). Mais identifier ce but défini en ces termes est loin de donner une théorie de l'explication scientifique. C'est lorsque l'on tente de circonscrire plus précisément ce but et de définir plus clairement ce qu'est une « explication satisfaisante » en sciences que les divisions se manifestent. L'élaboration d'une théorie de l'explication scientifique présuppose qu'il y ait une différence notable entre ce qui explique et ce qui n'explique pas, d'une part, et entre les explications dites scientifiques et les autres types d'explication, d'autre part. Et ce faisant, on retrouve un dilemme classique de la philosophie des

sciences, à savoir dans quelle mesure faut-il proposer des théories philosophiques reproduisant les traits de l'activité scientifique ?

Plusieurs modèles d'explication ont été proposés et les débats sont toujours actifs aujourd'hui. Toutefois, on accepte généralement certaines caractéristiques minimales, soit la cohérence interne, la cohérence externe et l'adéquation empirique. Mais on considère aussi que la notion d'explication est asymétrique, c'est-à-dire que si *A* explique *B*, *B* ne peut expliquer *A*. Dans ce qu'on peut appeler l'orthodoxie de l'explication scientifique, on retrouve le modèle déductivo-nomologique et le modèle causal-mécanique, soit la possibilité d'inférer l'explanandum à partir d'une loi ou d'une cause. Les critiques de cette orthodoxie ont surtout porté non pas sur la *pertinence* de l'approche, mais plutôt sur sa *nécessité* étant donné les exigences qu'elles placent sur l'explication. C'est-à-dire qu'un explanans exprimé d'après ces modèles serait suffisant mais non nécessaire à une « bonne » explication, à une explication satisfaisante. C'est pourquoi des modèles alternatifs sont non seulement envisageables mais aussi souhaitables en regard de certains objectifs louables visés par une explication. Il y a donc plusieurs manières de coordonner les efforts mobilisant les outils épistémiques, et parfois même ontologiques, engagés dans l'atteinte de l'un des objectifs associés à une explication scientifique satisfaisante : il y a donc plusieurs *stratégies explicatives*. Par exemple, l'explication en biologie, sans être radicalement différente de sa contrepartie physique, présente des particularités, comme un certain caractère historique. On peut aussi citer le modèle unificationniste et le modèle pragmatique. Enfin, ces différentes stratégies explicatives prennent différentes positions face au réductionnisme et au fictionnisme.

Ce pluralisme ne devrait toutefois pas cautionner un relativisme. Bien que plusieurs objectifs, parfois antagonistes, puissent être poursuivis avec une explication, tout n'explique pas pour autant. Il convient donc d'identifier clairement les éléments constitutifs de ce qui peut être légitimement considéré comme une bonne explication. Bref, de maintenir un pluralisme restreint. Au Chapitre 2 sera présenté un modèle d'explication faisant place à cette diversité sans s'y réduire, conservant ainsi un aspect normatif. Ce modèle s'articule autour de la notion générique de l'« expectabilité », comme quoi l'explanans fait en sorte que l'on soit en droit de s'attendre à l'occurrence du phénomène décrit par l'explanandum. Il est interprété selon deux critères, soit la modalité, selon laquelle l'explication doit montrer que ce qui s'est passé « devait » se produire, et la naturalité, où l'on montre, entre autres, que plusieurs situations similaires peuvent être subsumées sous un même concept. Afin d'illustrer l'opérationnalité de ce modèle, on se référera au concept d'espace des phases.

Ce modèle dresse un cadre explicatif et sert de paramètre aux discussions sur les concepts d'entropie et d'émergence. Car, tel que mentionné, la version thermodynamique de l'entropie est censée se

réduire à la version statistique parce que cette dernière l'expliquerait, d'une certaine façon qu'il reste à préciser, mais comme cette tentative de réduction est loin d'être convaincante, ou du moins certains doutes demeurent, on dit alors que cette propriété thermodynamique « émerge » de la dynamique moléculaire. Par ailleurs, et on pourrait dire à l'inverse, certains systèmes complexes, possédant certaines caractéristiques comme une structure dissipative, présentent une organisation ou forme particulière qui « émerge » suite à une dissipation d'entropie, autrement dit le principe entropique pourrait expliquer l'occurrence de certaines configurations structurales ou ce qu'on appelle l'auto-organisation.

1.2.3 Émergence

Dans son acception courante, le concept d'émergence renvoie justement à une « apparition soudaine », à ce qui paraît imprévu, inattendu, nouveau. Sans véritablement en donner une signification plus précise, ce concept se laisse aussi résumer par le slogan « le tout est supérieur à la somme des parties ». Le concept d'émergence a d'abord été élaboré, dans la littérature philosophique moderne, afin d'obtenir un compromis entre deux ontologies extrêmes mutuellement exclusives : une ontologie moniste et matérialiste, essentiellement réductionniste, où il n'y aurait que des éléments matériels et leurs propriétés, et une ontologie dualiste, donc antiréductionniste, où il y aurait la matière mais aussi l'esprit. Ainsi, bien qu'il n'y ait pas de consensus à son égard, le concept d'émergence est généralement reliée, de manière qu'il reste à préciser, au phénomène qui apparaît à partir de et dépend d'un phénomène plus *élémentaire* tout en étant dans une certaine mesure *autonome* de cette même base élémentaire. Le concept philosophique d'émergence est donc un concept essentiellement *relationnel* : les entités émergentes apparaissent ou surviennent d'entités plus fondamentales mais sont tout de même nouvelles ou irréductibles ou inexplicables par rapport à celles-ci. Il s'agit donc de concilier une certaine relation de dépendance, un « enracinement » (« *goundedness* »), avec une relative indépendance, une certaine autonomie. La relation entre ce qui émerge et à partir de quoi cela émerge étant souvent explicitée comme une relation entre un tout et ses parties, bref en termes *méréologiques*.

Il y a donc plusieurs questions épistémologiques et métaphysiques reliées à la notion d'émergence. En effet, il fait référence à des concepts tels que « imprévisibilité », « nouveauté », « irréductibilité », « inexplicabilité ». Mais peut-être que les phénomènes émergents n'existent pas vraiment, à l'instar d'épiphénomènes, c'est-à-dire de ce qui se surajoute à un phénomène sans exercer sur lui aucune réelle influence, bref des illusions ? Est-ce que la beauté d'un tableau, par exemple, est une propriété réelle ? Et le poids d'une table est-il émergent par rapport à celui de ses constituants ? Et puis, tout bien

considéré, s'agit-il seulement d'une dispute linguistique ? Si le concept d'émergence est pertinent, renvoie-t-il à une catégorie épistémologique ou ontologique ?

Puisque le concept d'émergence est compromissaire en ceci qu'il négocie entre deux extrêmes, l'un réductionniste et l'autre anti-réductionniste, la difficulté revient à caractériser les *relata*, que ce soit des propriétés, des concepts ou des prédicats, et la *relation* elle-même, avec une charge modale ou non, causale ou non, etc. Il s'agit donc de concilier, considérant ce qui précède, une relation de dépendance relative entre entités qui demeurent néanmoins relativement autonomes. Ce défi du schème dépendance-autonomie définit un continuum dans lequel s'inscrivent la plupart des thèses émergentistes. C'est pourquoi je présente une taxonomie de l'émergence (Chapitre 2). J'apporte également ma contribution au débat en définissant l'*émergence méthodologique*, qui exhorte à la recherche d'explication réductionniste autant que possible, sans multiplier les entités théoriques et sans pour autant coucher les phénomènes sur un lit de mystère, tombant par exemple dans un pluralisme exacerbé des niveaux d'organisation selon leurs différents modes d'accès, sans possibilité de les unifier.

Compte tenu du projet qui est le nôtre ici, soit celui de faire sens du concept d'entropie en ce qu'il se présente à la fois explanandum et explanans de l'émergence, l'accent sera mis sur l'analyse de l'articulation entre émergence et explication, c'est-à-dire de ce qui permet à l'émergence d'expliquer un phénomène et de ce qui permet d'expliquer un phénomène d'émergence, voire l'émergence en général. Adosser ainsi la notion d'émergence à celle d'explication n'a rien de nouveau et remonte à la conception empirico-logicienne des explications réductives.

1.2.4 Explanandum ou explanans ?

La motivation initiale de ce projet est ce constat – apparemment paradoxal – que l'émergence expliquerait l'entropie mais aussi que l'entropie expliquerait l'émergence. Tels quels ces deux énoncés ne peuvent être tous deux vrais. Autrement dit, ce qui explique semble aussi être expliqué de sorte qu'un concept, adjoint à des hypothèses particulières au sein d'un explanans, s'expliquerait lui-même. Le travail d'analyse et de désambiguïsation de cet apparent paradoxe doit alors faire face à la polysémie rampante des notions d'émergence, d'entropie et d'explication. Chacune de ces notions possède en effet sa propre taxonomie (Chapitres 2, 4 et 5). Nul doute donc qu'elles aient entraîné un lot d'assertions et de thèses les plus diverses et souvent contradictoires. La multitude de significations et d'interprétations de chacune d'entre elles prises individuellement déchaîne une multiplicité intenable de sens pouvant être accordés à leur combinaison au sein des deux énoncés précédents. Un paradoxe étant un énoncé sans solution, le présent paradoxe se dissout devant la solution consistant à assigner la bonne

signification au bon terme. Il faut alors faire preuve de prudence et apporter les nuances et les précisions qui s'imposent.

La solution formelle est donc la suivante : l'émergence (en un sens *A*) explique (en un sens *B*) l'entropie (en un sens *C*) et l'entropie (en un sens *D*) explique (en un sens *E*) l'émergence (en un sens *F*). Autrement dit, l'émergence peut (dans certaines conditions *X*) expliquer l'entropie, et l'entropie peut (dans certaines conditions *Y*) expliquer l'émergence. Il s'agit alors de préciser *A*, *B*, *C*, *D*, *E*, *F*, *X* et *Y*. Cette solution illustre le fait qu'il soit préférable de traiter ces deux questions, ces deux aspects du paradoxe, de front et non séparément.

En ce qui concerne l'entropie comme une *explanandum* de l'émergence, le débat s'articule autour de cette tentative de traduire le concept d'entropie thermodynamique au sein de la mécanique classique, qui, étant donné des contraintes épistémiques relativement évidentes, s'adjoint le concours des statistiques et des probabilités. Il s'agit donc d'une entreprise réductionniste et de son succès dépend largement l'exemplification du concept d'émergence (Chapitre 2). Évidemment, le succès ou l'échec de certains moyens épistémiques ne constituent pas un frein à diverses affirmations métaphysiques. Et le concept d'émergence n'y fait pas exception. Il s'agit plutôt de justifier le cas échéant ce genre d'affirmations. Mais il faut aussi établir les conditions de succès d'une réduction et ce que cela implique pour le concept d'émergence. Or, affirmer un cas d'émergence revient déjà à caractériser une situation théorique devant ces efforts réductionnistes. Ce succès et ces efforts sont jugés à l'aune d'une théorie de l'explication car généralement une réduction implique une explication.

En ce qui concerne l'entropie comme un *explanans* de l'émergence, deux approches sont ici distinguées, soit l'*approche substantielle* et l'*approche analogique*. La première approche renvoie à la stratégie généralisée du principe entropique (Chapitre 4), où, en termes simples, un système absorbe de l'énergie et rejette de l'entropie afin de former une certaine organisation structurale. Devant le constat de cette organisation (qu'on appelle aussi « ordre », « complexité », etc.), comme ce qui est observé dans le domaine du vivant, plusieurs y ont vu une violation du principe entropique interprété comme un accroissement du désordre. Au contraire, le phénomène de la vie est désormais considéré comme un allié important du principe entropique et du second principe de la thermodynamique, puisque les composés chimiques biotiques génèrent beaucoup plus d'entropie qu'ils ne le feraient dans un état abiotique (Capek & Sheehan 2005 : 308). Une interprétation et une application rigoureuse du principe entropique permettent donc de voir que cette organisation ne constitue ni une violation de ce principe, ni une instanciation de ces événements improbables qu'autorise sa version statistique.

L'approche analogique s'appuie sur une ou des analogie(s) formelle(s) entre la structure ou des propriétés mathématiques de l'entropie en physique statistique et celle de la formalisation des processus

biologiques évolutionnistes. La stratégie commune de cette approche est d'identifier une population de sous-unités d'un large système comme éléments d'un espace des phases servant à définir un concept spécifique d'entropie et d'y attribuer la propriété de non-décroissance (ou plus simplement d'augmentation) de l'entropie. Par exemple, Fisher (1922 et 1930) a comparé une population de facteurs mendéliens (gènes) à un ensemble de molécules constituant un gaz où la multitude des causes individuelles peut être négligée et l'attention portée sur le comportement statistique de l'ensemble. En tout état de cause, si cette stratégie est pour être effective et calquer la description des systèmes qu'elle s'emploie à expliquer sur celle de l'entropie statistique, encore faut-il que l'entropie possède les caractéristiques recherchées, comme la non-décroissance, ce qui n'est pas assuré.

On constate donc que le concept d'entropie est mobilisé au sein de stratégies explicatives diverses mais néanmoins fort ambitieuses en ceci qu'elles traversent généralement plusieurs disciplines. Malgré une divergence dans les objets mais aussi dans les concepts et les méthodes, on constate aussi une certaine convergence dans les stratégies explicatives. La récupération du formalisme de la mécanique statistique pour l'explication de phénomènes biologiques en est un exemple. C'est aussi souvent le cas du traitement des systèmes complexes exigeant des approches multi-, inter- et transdisciplinaires.

1.3 Plan de la thèse

Le prochain chapitre (Chapitre 2) situe le contexte de la philosophie des systèmes complexes en présentant et en prenant position sur plusieurs enjeux de la philosophie méthodologique (ou générale) des sciences, en particulier sur les débats entourant les notions (aussi polysémiques) d'explication et d'émergence (j'ai laissé de côté l'historique de ce concept). Plus précisément, il propose un cadre réaliste à cette analyse, offre un modèle pluraliste (restreint) de l'explication, définit un concept d'émergence méthodologique et un cadre d'analyse de l'« émergence expliquée » et de l'« émergence qui explique ».

Le Chapitre 3 présente l'historique du concept d'entropie, principalement dans ses versions thermodynamique, avec les travaux de Carnot, Kelvin et Clausius, et statistique, avec ceux de Maxwell, Boltzmann et Gibbs. Un aperçu de ses récupérations en mécanique quantique, en génétique et en théorie de l'information est ensuite présenté.

L'examen critique du concept d'entropie commence au Chapitre 4, avec une discussion approfondie de sa version thermodynamique. Celle-ci débute par une analyse de l'irréversibilité, placée dans le contexte de la (vaste) question philosophique du temps, et en particulier des asymétries temporelles. J'analyse la signification du second principe de la thermodynamique, défini en termes de chaleur et de travail, et qui est donc distingué du principe entropique, défini à partir de l'entropie. Le chapitre se termine par une analyse de la stratégie explicative de ce principe entropique exemplifiant l'irréversibilité.

Le Chapitre 5 poursuit l'examen critique du concept d'entropie, dans sa version statistique cette fois. Une discussion sur la notion de probabilité semble alors inévitable. J'offre une interprétation originale des probabilités en insistant sur la méthode d'attribution empirique de valeurs de probabilités. Quatre versions différentes du concept statistique d'entropie (entropies continue, combinatoire et macroscopique de Boltzmann et entropie de Gibbs) ainsi que cinq interprétations différentes (probabilité, désordre, information, travail et dispersion) sont ensuite discutées.

Enfin, le Chapitre 6 dissout ce paradoxe de la symétrie de l'explication de l'entropie et de l'émergence en montrant sous quelles conditions l'entropie peut être explanandum, dans un premier temps, et explanans de l'émergence, dans un second temps, selon deux approches, soit l'approche substantielle et l'approche analogique.

J'ai tenté d'emprunter un style le plus clair possible et d'être aussi précis que possible. S'il peut y avoir quelques répétitions, c'est uniquement dans le but de faciliter la lecture, au demeurant complexe.

2 Émergence et explication

RÉSUMÉ. Les niveaux d'organisation, comme celui des particules subatomiques ou des organismes vivants, peuvent être décrits selon une approche réductionniste et moniste où il n'y aurait qu'un seul niveau fondamental, ou plutôt selon une approche pluraliste visant à les dissocier en insistant sur ce qui les différencie. Le concept d'émergence, qui tente de concilier dépendance et autonomie entre niveaux, a été développé dans l'idée d'offrir une position médiane entre ces deux positions extrêmes. La polysémie de ce concept se rapporte autant à ses relata (propriétés, substances, prédicats) qu'à la relation qu'il définit (ontologique ou épistémologique, synchronique ou diachronique). Mais en cherchant à caractériser cette relation, il semble impossible de faire l'économie de l'analyse des théories scientifiques décrivant ces niveaux et donc du concept d'explication. Ce concept, tout aussi polysémique, qui est ici d'abord présenté selon l'orthodoxie (lois et causes) et ses alternatives, est ensuite défendu selon un pluralisme restreint par deux critères, soit la modalité et la naturalité. Les conditions pour que l'émergence soit expliquée, mais aussi pour que l'émergence puisse expliquer, sont enfin discutées, où je propose l'émergence méthodologique, compatible avec des explications de type réductif.

*Confusion is a word we have invented
for an order which is not understood.*

Henry Miller

Que la réalité se présente par niveaux semble un fait difficilement discutable : des atomes forment des molécules, lesquelles forment des cellules, qui forment des organismes, lesquels peuvent former des populations et des sociétés. Que ce soit discutable ou non, cependant, deux tentations guettent l'intrépide qui désire en rendre compte : d'abord, accepter, justement comme un fait, ces constatations, sans doute pour des raisons d'évidence, et tenter de définir ce qu'est un « niveau de réalité » et leurs relations mutuelles, ou encore, sans doute pour des raisons d'unité, vouloir ramener ces différents niveaux à un seul en traduisant leur description dans les termes propres à ce niveau jugé plus fondamental. Là commencent les difficultés.

De toute évidence, déterminer en quoi consiste la réalité, qu'elle soit par niveaux ou non, et pour autant que l'on s'entende sur ce que cela signifie, n'est pas une mince affaire. Ce n'est pas pour rien que plusieurs philosophes ont été amenés à nier la possibilité même de le faire. Cette possibilité, ou non, de

déterminer en quoi consiste la réalité renvoie aussitôt au *problème général de la connaissance*, lequel peut être exprimé selon trois aspects (Tournier 1988 ; Bunge 2006) : (i) qu'est-ce que la connaissance et quelle est sa valeur (aspect épistémologique) ? (ii) quelle est la nature des objets de connaissance (aspect ontologique) ? (iii) quels sont les mécanismes de production de connaissance (aspect psychologique) ? C'est ce qui a alimenté le débat entre l'*empirisme*, et ses avatars, et le *réalisme*, principalement dans sa version dite scientifique (voir Chakravartty 2007 ; van Fraassen 2008). Les termes de ce débat se sont généralement concentrés autour de ce qui peut être qualifié de *justification* pour une croyance donnée, de ce qui est en fait un discours « valide » ou « fondé épistémologiquement »¹.

Ainsi, l'empirisme évacue du domaine de la justification ce qui n'est pas « empiriquement accessible » ou « observable », et se montre très circonspect envers un discours à portée ontologique ou métaphysique, c'est-à-dire portant sur la nature des choses ou de la réalité, particulièrement si elle est inobservable. Le réalisme prétend au contraire pouvoir fournir des justifications à ce qui n'est pas « empiriquement accessible » ou « observable », que cela puisse même servir de base à une explication, et qu'un discours prétendant décrire la réalité puisse être valide dans une certaine limite qui reste à définir. C'est pourquoi, selon Salmon (2005), le travail de réconciliation entre ces deux thèses revient à savoir si oui ou non le raisonnement inductif possède les ressources suffisantes pour obtenir des preuves (observationnelles) pour ou contre un discours, ou des énoncés à propos d'entités, propriétés, états, etc. inobservables. Autrement dit, est-il légitime d'inférer, à partir d'observations à la base d'une induction, des conclusions sur ce qui n'est pas observable ?

Ces considérations situent bien entendu la discussion dans un large contexte épistémologique. Il est cependant important de la situer ainsi puisque toute discussion portant sur les niveaux, des atomes aux sociétés, implique un ensemble de problèmes et de débats importants et persistants en philosophie des sciences. On peut par exemple avancer que les propriétés *en apparence* différentes d'un système composé (tout) ne sont pas *en réalité* différentes de celles de ses composants (parties). Que ses propriétés soient en définitive, tout bien considéré, différentes de celles des composants est généralement inféré de l'impossibilité d'en donner une *explication* (section 2.1) en termes de lois régissant les propriétés des composants. Par exemple, il ne semble pas possible d'expliquer les phénomènes biologiques uniquement au moyen des lois de la physique, ou encore il n'est pas évident que les propriétés thermiques de larges systèmes puissent être expliquées ou inférées des propriétés mécaniques des molécules qui les composent. Mais déjà la notion d'explication dresse des obstacles car plusieurs

¹ En toute rigueur, l'étude de la connaissance en générale est l'objet de la gnoscologie (en anglais, « *theory of knowledge* » ou « *epistemology* »), tandis que l'étude de la connaissance scientifique est celui de l'épistémologie (en anglais, « *philosophy of science* »). (Voir Jodoin 2014a.)

objectifs peuvent être poursuivis avec une explication : comprendre, prédire, contrôler. Il n'est pas évident qu'ils soient tous compatibles ou qu'ils puissent être satisfaits par un même concept d'explication. S'il est entendu que des lois peuvent jouer un rôle explicatif, il en va de même des causes. Il faut donc définir clairement ce qu'est une loi ou une cause, ce qui est voué à la polémique. De plus, plusieurs explications scientifiques jugées satisfaisantes se passent de loi ou de cause, ou du moins elles en font un usage différent très indirecte, différent de ce que plaide l'orthodoxie. De fait, plusieurs modèles d'explication alternatifs ont été proposés et sont aussi utilisés. Il faut donc faire place à un certain pluralisme de l'explication sans toutefois tomber dans un relativisme susceptible d'obscurcir cette notion, primordiale en sciences comme en philosophie, plutôt que de l'éclairer.

Tout compte fait, si l'on dispose d'une conception satisfaisante d'une explication scientifique, l'impossibilité même, en pratique ou en principe, d'une explication semble impliquer des conséquences importantes pour l'étude critique des sciences. Ainsi, l'échec d'une explication (exhaustive) des phénomènes biologiques en termes physiques semble bien être une bonne raison pour laquelle la biologie et la physique forment des disciplines distinctes. Mais s'il était possible de ramener les lois ou les concepts biologiques à ceux de la physique, alors serait opéré ce qu'on appelle une *réduction* (section 2.3). Les critères visant à déterminer si une certaine réduction est réussie exigent d'abord que soit précisé ce qui doit être réduit, génériquement, comme des propriétés, des concepts ou des lois, et spécifiquement, au sein de théories particulières, comme les espèces, les organismes ou les gènes dans le cas de la biologie. Par exemple, dans la préface de l'édition pour le trentième anniversaire de son fameux livre, *The Selfish Gene*, Dawkins (2006) pose la question de l'identification du niveau où opérerait la sélection naturelle, soit le « niveau égoïste », et propose évidemment le gène. Comme il identifie les populations d'organismes issues de cette sélection comme celles maximisant les copies de leurs gènes, et propose ainsi d'expliquer cette sélection par l'action de ces gènes, il en vient donc à réduire (prétendument) l'ensemble de la biologie à la génétique. En sorte que les critères d'une réduction renvoient généralement à des considérations épistémologiques comme, une fois de plus, l'explication, ou encore la déduction ou la prédiction. Il ne va donc pas de soi qu'un discours sur différents niveaux d'un système et de ses parties porte bien sur leur réalité, qu'il puisse en révéler la vraie nature, qu'il puisse même fournir des outils rationnels pouvant guider nos comportements nonobstant sa vérité.

Advenant l'échec d'une telle explication, de type réductif, les propriétés du système au niveau supérieur, c'est-à-dire du système composé, sont souvent considérées comme exemplifiant le concept d'*émergence* (section 2.3). Ce concept entretient donc un lien étroit avec celui de réduction. Mais déjà on peut concevoir que ces concepts peuvent recevoir plusieurs significations différentes selon les termes utilisés pour les définir. C'est-à-dire que les énoncés « *X* se réduit à *Y* » ou « *X* émerge de *Y* » renvoient

à des concepts différents en fonction de ce à quoi réfèrent X et \mathcal{I} , que ce soient (tel qu'évoqué) des propriétés, ou encore des lois, des concepts ou des prédicats, mais aussi en fonction de leur relation.

Ces considérations épistémologiques définissent un cadre d'étude des *systèmes complexes* (Chapitre 6), c'est-à-dire des systèmes constitués d'un grand nombre d'éléments interagissant de manière non triviale ou non simple, selon la conception classique de Simon (1962). En effet, ces systèmes, point focal de plusieurs champs de recherche, peuvent être étudiés à plusieurs niveaux, de complexité variable et faisant l'objet de diverses théories scientifiques, de la géographie physique à la socio-économie. Ces systèmes sont généralement associés à des notions telles que la non-linéarité, la non-additivité, l'irréversibilité, la brisure de symétrie, l'ordre, l'émergence, la hiérarchie, l'historicité, etc. (Hooker 2011) Les systèmes complexes présentent ainsi, généralement, des composantes qui sont, à un certain degré ou à un autre, indépendantes, et donc autonomes dans leur comportement, tout en subissant diverses interactions ; il en résulte un comportement global du système difficilement prévisible ou explicable sur la base de la connaissance de ses composants.

L'analyse de tels systèmes se penchent sur ce qui est donc appelé des *niveaux d'organisation* (section 2.2) présentant des caractéristiques suffisamment stables pour être étudiées, mais qui peuvent néanmoins être « entremêlés » et donc présenter des obstacles à cette analyse. Celle-ci, tel que mentionné, renvoie inévitablement à une panoplie de concepts visant à caractériser ces niveaux et demande aussi d'explicitier le point de vue épistémologique adopté. L'approche à l'œuvre ici découle de la thèse générique du réalisme scientifique, sans pour autant endosser la croyance que les théories scientifiques puissent offrir le « fin mot de l'histoire » sur l'existence des entités naturelles, réelles. Les thèses de Wimsatt et Bunge seront ensuite utilisées afin d'encadrer une conceptualisation des niveaux d'organisation.

Ces caractéristiques des systèmes complexes constituent une nouveauté à laquelle font face diverses disciplines scientifiques et qui a su déstabiliser l'approche réductionniste longtemps dominante. C'est pourquoi la notion d'émergence fait à l'heure actuelle l'objet de plusieurs discussions philosophiques. Cette nouveauté, voire ce nouveau paradigme pour certains (par ex. Prigogine & Stengers 1979), est tributaire d'une vaste révolution épistémologique débutée au XIX^e siècle avec la thermodynamique classique et qui s'est poursuivie durant la seconde moitié du XX^e siècle avec la thermodynamique du non-équilibre (Chapitre 3).

Le concept d'entropie est issu de cette révolution épistémologique. Ce concept, émanant d'abord de la thermodynamique (Chapitre 4) puis reformulé en mécanique statistique (Chapitre 5), apparaît particulièrement intéressant pour l'étude de la relation entre niveaux d'organisation de systèmes complexes (Chapitre 6). En effet, la mécanique statistique est une théorie tentant d'expliquer le

comportement observable (niveau macroscopique) de grands systèmes par la dynamique de molécules inobservables (niveau microscopique). L'entreprise visant à faire sens du concept d'entropie et des prétentions explicatives le mobilisant est donc aux prises avec un réseau de problèmes épistémologiques de taille, allant aux fondements même de la connaissance (évoqués plus tôt).

On peut résumer (si faire se peut) le problème philosophique qui nous concerne ici particulièrement de la façon suivante. Le concept d'entropie est polysémique car il apparaît dans (au moins) deux théories différentes, la thermodynamique et la mécanique statistique. Dès lors se pose la question du statut respectif de ce concept au sein de ces théories (que signifient-ils ?) et celle de leur relation (y en a-t-il un meilleur que l'autre ?). La réponse à la première question exige un terrain d'analyse fiable, un point de vue justifié, car les hypothèses à la base des diverses définitions de l'entropie sont fort différentes. Or, il est loin d'être évident que l'approche de la mécanique statistique, basée sur l'atomisme et plus abstraite en ce qu'elle décrit des phénomènes inobservables, soit plus fondamentale ou plus justifiée que celle de la thermodynamique, que l'on qualifie de « phénoménologique » et dont les lois ont été maintes fois corroborées, mais qui conçoit la matière comme une « boîte noire » car elle est muette quant à sa structure. La seconde question est souvent abordée par le spectre du débat (déjà évoqué) à savoir si l'un de ces concepts se réduit à l'autre et que signifie cette réduction.

Puisqu'un concept se réduirait à un autre si le premier expliquait d'une certaine manière le second, il faut expliciter la notion d'explication, mais aussi celle d'émergence car celle-ci est souvent considérée comme exemplifiée dans plusieurs cas d'échec de la réduction. L'émergence méthodologique, que je défends sans pourtant soutenir qu'elle devrait remplacer tous les concepts d'émergence existants, affirme la possibilité de certaines explications de type réductif et tente d'éviter de réduire la relation d'émergence à un mystère ou à un effet de surprise.

2.1 Explication

Les débats concernant la nature et la forme de l'explication scientifique n'ont pas perdu de leur intensité et il n'est pas question ici de tenter d'y mettre un terme. Puisqu'il n'y a pas de consensus quant à ce que devrait accomplir une explication scientifique, il y a lieu d'être dubitatif quant à une éventuelle résolution de ces débats. Au sein des discussions de ce qu'on appelle la philosophie méthodologique (ou générale) des sciences réside en effet ce que j'appellerais le *dilemme de la censure* : doit-on, et si oui dans quelle mesure, proposer des théories philosophiques reproduisant les traits de l'activité scientifique ? Ainsi, une éventuelle contradiction entre une théorie de l'explication et ce qui se fait en pratique, en général ou au sein d'une certaine communauté de scientifiques, constitue-t-elle une raison suffisante, ou simplement une bonne raison, pour réfuter une théorie philosophique de l'explication scientifique ? Une telle théorie se place donc sur le mode normatif puisqu'elle propose une censure, même partielle, du discours scientifique. Son adepte se voit éventuellement contraint d'entraver une activité scientifique, même pour de bonnes raisons épistémologiques, susceptibles de donner des résultats pragmatiques importants. Un philosophe risquerait-il, par exemple, de condamner une théorie oncologique importante, pouvant apporter un remède contre le cancer, sous prétexte que celle-ci serait incompatible avec sa théorie fétiche, car elle mettrait à contribution, par exemple, des causes et non pas des lois ?²

Ce dilemme de la censure est à l'origine, je crois, de bien des circonspections et circonvolutions en philosophie des sciences. En tentant une théorie unifiée de l'explication scientifique, il devient pratiquement impossible de rendre compte de la diversité des explications en sciences. Une théorie trop générale risque la trivialité ou l'incompatibilité avec cette diversité, tandis qu'une théorie voulant en rendre compte de trop près risque d'être purement descriptive, sans pouvoir d'unification et sans intérêt proprement philosophique. Or, ces difficultés, ces risques guettent autant le philosophe portant un regard qui se veut critique sur le discours scientifique que le scientifique voulant rendre compte du monde. Si, pour ce faire, il est patent que le scientifique cherche à élaborer des théories, il est moins évident d'identifier ce qu'il cherche vraiment par là : quelle genre de théorie, et une théorie pour quoi ? Il cherche peut-être une théorie ayant telle ou telle caractéristique dans le but d'expliquer, de comprendre, de décrire, d'unifier, de prédire ou même de contrôler un phénomène particulier. Ces buts ne sont pas identiques, peuvent difficilement se ramener à un seul et il ne va de soi que l'un d'entre eux soit plus louable ou souhaitable qu'un autre.

² Parlant de ces « reconstructions rationnelles », comme les théories de l'explication scientifique, Batterman (2002 : 3) semble rejeter ce caractère normatif : « *All too often, however, these reconstructions end up being quite far removed from the actual science being done.* »

L'élaboration d'une théorie de l'explication scientifique présuppose qu'il y a une différence notable entre ce qui explique et ce qui n'explique pas (comme une simple description), d'une part, et entre les explications dites scientifiques et les autres types d'explication, d'autre part. Au sens le plus large, une explication, on l'a dit, qu'elle soit scientifique ou non, est censée réaliser plusieurs tâches : montrer que ce qui s'est produit devait se produire ou qu'on était en droit de s'y attendre, faire comprendre, d'unifier plusieurs phénomènes semblables, permettre de prédire, de contrôler, de manipuler. Bien que cette liste ne soit pas exhaustive, on constate qu'il y a plusieurs recoupements entre ce que l'on cherche d'une explication et ce que cherche à obtenir le scientifique en général, ce qu'il poursuit ou juge avoir une certaine valeur dans son travail. Avec une portée aussi large, il devient alors presque consensuel de dire que le « but de la science » consiste à générer des « explications satisfaisantes » (Popper 1991 : 297). Mais identifier ce but défini en ces termes est loin de donner une théorie de l'explication scientifique. C'est lorsque l'on tente de circonscrire plus précisément ce but et de définir plus clairement ce qu'est une « explication satisfaisante » en sciences que les divisions et les problèmes se manifestent. Par exemple, il peut y avoir d'excellentes raisons épistémologiques d'écarter une explication *ad hoc* même si celle-ci peut avoir des vertus pragmatiques importantes, comme celles de permettre de prédire et de contrôler un phénomène particulier³. C'est l'une des raisons pour lesquelles on peut facilement trouver un fossé considérable entre le modèle et la pratique, exemplifiant ainsi le dilemme de la censure.

Devant les difficultés d'une théorie unifiée de l'explication scientifique, certains ont récusé ce but et avec lui le caractère normatif de l'explication en sciences, celles-ci se limiteraient, selon eux, à décrire, représenter, prédire. Comme si une description, une représentation, ou une prédiction était nécessairement dépourvue de normativité, comme si elle ne pouvait être bonne ou moins bonne, qu'il n'y en aurait pas de moins scientifique que d'autres. Rejeter toute normativité du discours philosophique sur les sciences revient en définitive à le cantonner à des comptes-rendus. Ce caractère normatif apparaît donc important dans le cadre d'un critère de démarcation du discours scientifique des autres discours plus ou moins rationnels, comme celui du sens commun ou de la poésie, en bref d'un critère de scientificité. Le réputé critère de falsifiabilité (Popper 1935) peut sembler nécessaire, encore qu'on puisse y objecter d'être éloigné de ce que nous dit l'histoire des sciences (Kuhn 1962). Mais il ne permet pas de circonscrire univoquement le discours scientifique. Il n'est plus question, aujourd'hui, de discriminer un discours sur des considérations ontologiques, possiblement en raison des

³ Une théorie permettant de faire des prédictions précises quant au comportement d'un système physique ou économique pourrait être considérée comme « explicative » même si ses prédictions ne seraient pas avérées, par exemple, les mercredis. Et on peut se douter des réticences à insérer cette théorie au sein d'une conception métaphysique du monde où les mercredis auraient un statut ontologique particulier !

profonds bouleversements que les développements scientifiques ont fait subir à l'image que l'on se faisait du monde et qu'en conséquence l'influence s'orienterait davantage de la science vers l'ontologie plutôt que l'inverse. Le critère explicatif peut jouer ce rôle normatif et définir le discours scientifique par ce qui est falsifiable *et* ce qui explique est déjà plus restrictif. Il pourrait s'agir d'un autre critère normatif mais celui de l'explication semble plus fort, d'autant plus si l'on endosse la définition poppérienne du but de la science. Malgré les défis de théorisation qu'elle représente, la notion d'explication est centrale en sciences comme dans l'examen critique qu'en propose la philosophie des sciences.

Toute théorie de l'explication scientifique présente idéalement deux classes de *pertinence*, l'une concernant les outils épistémologiques privilégiés, comme, par exemple, les lois ou les causes, et l'autre circonscrivant, pour un contexte théorique donné, les données, les variables pertinentes à ces outils épistémologiques pour que l'explication puisse jouer son rôle. Pour une théorie scientifique donnée, tout ne peut pas faire l'objet d'une explication, autrement dit tout ne peut pas être une « question-pourquoi » légitime (van Fraassen 1980 : 126). Dans la théorie newtonienne par exemple, on évite les questions telles que « pourquoi les corps maintiennent-ils leur vitesse en l'absence de forces ? » (car on ne saurait accepter une explication circulaire en stipulant le principe d'inertie), ou encore « pourquoi y a-t-il la gravité ? » (à moins d'accepter quelque chose comme « parce qu'il y a des masses »⁴). Il y a donc ce qu'on peut appeler des « faits bruts » qui ne font pas l'objet d'explication pour une théorie donnée. Ceux-ci peuvent être définis, mais il est entendu qu'une définition ou une stipulation n'explique pas. Parmi la somme colossale de données ou d'information pouvant être amassées pour un système donné, lesquelles doivent être retenues, lesquelles sont pertinentes à l'explication dont le but serait de faire comprendre, de prédire, de contrôler, ou de manipuler ? Par exemple, pourquoi la couleur de la boule de billard est-elle impertinente à l'explication de son déplacement ? Il faut donc savoir écarter le « bruit explicatif » (Batterman 2002 : 4). En général, les discussions philosophiques traitent principalement de la première classe de pertinence, tandis que les théories scientifiques sont censées circonscrire la seconde.

En plus de ces deux classes de pertinence, une explication scientifique doit respecter certaines caractéristiques minimales, soit la *cohérence interne*, la *cohérence externe* et l'*adéquation empirique*. La cohérence interne exige minimalement que l'explication ne possède pas d'éléments contradictoires, que la théorie qui la supporte soit consistante. Il y a, de plus, une autre condition formelle, qui est celle de l'*asymétrie*. C'est-à-dire que si *A* explique *B*, *B* ne peut expliquer *A*. Le critère de la *cohérence externe* étend cette

⁴ On se rappellera le « *hypothesis non fingo* » de Newton.

exigence à d'autres théories jugées valides, mais il n'est pas aussi fort que le premier. Par exemple, on s'attend généralement à ce qu'une théorie génétique soit compatible avec la mécanique quantique. Dans le cas contraire, une justification est nécessaire, mais la nouvelle théorie contredisant, même partiellement, une théorie déjà établie ne sera pas nécessairement rejetée. Le troisième critère est relativement fort puisqu'une explication incompatible avec les données empiriques ne peut être considérée acceptable ou correcte, ou alors ce sont ces données empiriques qui doivent être revues. Mais ce critère peut même être renforcé en exigeant que l'adéquation empirique se rapporte à la *testabilité*, c'est-à-dire que non seulement une hypothèse explicative doit être en accord avec les faits observés, mais elle doit l'être aussi avec les nouveaux faits, non encore observés, qu'elle projette⁵. Même ces caractéristiques ne s'appliquent pas aussi aisément qu'on le voudrait. Car une explication d'un phénomène particulier exige plus souvent qu'autrement non pas une mais plusieurs théories scientifiques dont la conjonction peut mener à des difficultés d'interprétation de certains de leurs concepts. De plus, on veut bien une explication qui ne contredise pas les faits, mais il est fréquent qu'une explication ait recours à des idéalizations qui sont, *stricto sensu*, fausses (section 2.1.1.2).

Malgré les difficultés de définir, *lato sensu*, une explication, surtout dans son caractère normatif, son rôle définit plusieurs autres notions en philosophie des sciences. Et le résultat peut paraître parfois circulaire. Ainsi, une réduction épistémologique, où les concepts d'une théorie sont « ramenés » à ceux d'une autre jugée supérieure (section 2.1.3), peut être considérée comme valide si ces concepts sont d'une certaine façon « expliqués » par cette théorie supérieure. Corrélativement, un phénomène peut être considéré « émergent » s'il n'est pas possible d'en fournir une explication de type réductif. Au surplus, une théorie est jugée scientifique ou simplement valide si elle est en mesure d'expliquer, alors même qu'on juge ce qui explique et ce qui est expliqué à l'aune d'une théorie particulière. La notion d'explication est donc difficile à caractériser clairement puisqu'elle est somme toute primitive et qu'elle intervient elle-même dans la caractérisation d'autres notions importantes, comme les notions de réduction, d'émergence et de théorie scientifique.

Il y a donc plusieurs difficultés à offrir une théorie normative, prescriptive et proscriptive, de l'explication. Il faut en effet déterminer l'objectif rattaché à cette normativité. Or il va de soi que nos intérêts, comme ceux des scientifiques, sont très divers. En sorte qu'ils orientent ce qui est visé par une explication. Toutefois, tel que discuté avec le dilemme de la censure, il est loin d'aller de soi que l'activité, *de facto*, des scientifiques implique, *de jure*, ce que doit être une explication, pas plus qu'elle

⁵ Cette exigence se rapproche toutefois de celle de la cohérence externe : « *testability, is supposed to exclude non-scientific explanations that make reference to explanatory factors that cannot be subject to confirmation or disconfirmation by observation, experiment or other empirical data.* » (Rosenberg 2001 : 31)

puisse réfuter une théorie particulière de l'explication. De plus, il n'est pas plus évident que l'atteinte d'un de ces objectifs soit une garantie de la vérité de l'explanans ou qu'il soit en mesure de satisfaire les desiderata d'une explication. L'échec, même contingent, d'une théorie unifiée de l'explication n'est pas nécessairement un obstacle à la réalisation de l'un des divers objectifs poursuivis avec une explication scientifique, comme ceux visant à faire comprendre ou à procurer un certain contrôle sur les phénomènes. Face à ce pluralisme d'objectifs poursuivis ou connotés par la notion d'explication, il va de soi que certaines explications peuvent paraître « meilleures » que d'autres en regard du ou des critères adoptés, qui sont eux-mêmes conditionnés par les objectifs poursuivis. Ainsi, une explication jugée moins « bonne » selon un certain critère peut tout de même satisfaire certains des objectifs discutés précédemment. Par exemple, on peut « expliquer » certains phénomènes thermiques par la loi de Fourier, comme quoi, en termes simples, un gradient de température induit un flux de chaleur, de manière à offrir les conditions pour un certain contrôle sur ces phénomènes ; mais on pourra objecter que cette explication ne fournit pas une « compréhension fondamentale » de ces phénomènes, ou encore qu'elle ne permet pas une unification de ceux-ci avec des phénomènes à plus petite échelle, celle des molécules par exemple. À l'inverse, une explication faisant appel au mouvement des molécules pour expliquer ces phénomènes thermiques pourrait bien se faire objecter qu'elle ne procure pas ces « conditions de contrôle ».

Cette discussion sur les difficultés de théoriser l'explication vise à encadrer le débat sur les vertus épistémologiques de certains concepts et principes scientifiques – au premier chef le concept d'entropie et le principe entropique. Car ces vertus sont souvent basées sur l'explication. Dans ce cadre d'analyse, il pourrait sembler avisé d'identifier une théorie de l'explication et d'en tirer les conclusions philosophiques pertinentes. Mais le résultat ne serait que partiel car limité à cette théorie particulière. La discussion qui suit placera en exergue la pluralité des stratégies visant l'atteinte de ces multiples objectifs. Il y a donc plusieurs *stratégies explicatives*, c'est-à-dire qu'il y a plusieurs manières de coordonner les efforts mobilisant les outils épistémiques, et parfois même ontologiques, engagés dans l'atteinte de l'un des objectifs associés à une explication scientifique satisfaisante. Cependant, cela ne signifie pas que toutes ces stratégies soient équivalentes et que cette pluralité implique un pluralisme éliminant par le fait même le caractère normatif de l'explication. Tout n'explique pas. Les prochaines sections tenteront de circonscrire davantage, au-delà des caractéristiques minimales discutées précédemment, ce que doit être une explication.

Dans ce qui suit je propose de présenter plusieurs tentatives de théorisation de l'explication scientifique en séparant ce que j'appelle l'orthodoxie de ces théorisations et ce qui ne l'est pas, ses alternatives.

2.1.1 Orthodoxie : lois et causes

L'orthodoxie de l'explication scientifique renvoie d'abord au *modèle déductivo-nomologique*, qui tire son nom de « déduction » (*deductio*) et de « loi » (*nomos*). Il consiste essentiellement à subsumer ce qui doit être expliqué, l'*explanandum*, sous ce qui est censé expliquer, l'*explanans*. Il s'agit de déduire (rétrospectivement) un phénomène particulier de la conjonction d'énoncés généraux, en particulier des lois, et de conditions initiales. Il existe donc ici une connexion forte entre explication et prédiction. Bien entendu, ce modèle possède ses forces et ses faiblesses, lesquels ont donné naissance à des critiques et à des alternatives notoires (pour une vue d'ensemble, voir Woodward 2002 ; 2003a ; 2003b). Cette orthodoxie comporte un autre modèle, complémentaire au modèle déductivo-nomologique, en certains points, en ce qu'il tente de répondre à ses critiques, soit le *modèle causal-mécanique* où, pour l'essentiel, une cause joue le rôle principal dans l'*explanans*. Au cœur de cette orthodoxie se trouve donc la possibilité d'inférer l'*explanandum* à partir d'une loi ou d'une cause⁶.

Le modèle déductivo-nomologique, aussi appelé le « modèle des lois subsumantes » (*covering-law model*), demande ainsi la caractérisation de ce qu'est une loi (voir Hempel 1965 ; Armstrong 1983 ; Carroll 2010). Il faut alors distinguer plusieurs choses ici. D'une part, un énoncé issu d'une communauté scientifique portant le nom de « loi » n'implique pas la satisfaction de critères formels et épistémologiques jugés nécessaires, et le cas échéant, il n'est pas dit qu'il y ait une détermination ontologique réellement à l'œuvre, qu'il existe ce qu'on appelle des « lois de la nature ». Il faut donc distinguer le projet d'identification de régularités ou de corrélations entre propriétés ou variables du projet de leur interprétation. D'autre part, il est ardu d'identifier ces critères formels et épistémologiques, bref des critères de nomicité, car si l'on accepte qu'une loi soit une régularité universelle, sans exception, alors on fait face au problème des « généralisations accidentelles », soit des régularités universelles qui ne sont pas des lois. Par exemple, « toute sphère d'uranium a un diamètre inférieur à 100 000 km » peut être considéré comme une loi sur la base de théories physiques stipulant l'instabilité critique d'une telle structure, mais « toute sphère d'or a un diamètre inférieur à 100 000 km » serait plutôt considéré comme une généralisation accidentelle, un fait contingent néanmoins exemplifié partout. Malgré ces difficultés, certains critères de nomicité ont été proposés : que l'énoncé soit universel, qu'il soit suffisamment corroboré par l'expérience, qu'il supporte les contrefactuels, qu'il puisse jouer un rôle unificateur au sein d'une théorie. Aucun de ces critères ne fait toutefois consensus. La question se pose donc à savoir si ce modèle peut être effectif même s'il n'y a pas de définition claire de ce qu'est une loi, de ce qui fait qu'une généralisation puisse être une loi.

⁶ Il est entendu que cette orthodoxie n'a rien d'officielle mais vise plutôt à circonscrire les thèses dominantes.

Ensuite, il est possible que la régularité servant à expliquer un phénomène ne soit pas stricte mais de nature statistique ou probabiliste, auquel cas elle perd en universalité et l'explication perd donc de sa force, sans être pour autant dénuée de valeur. Hempel lui-même a proposé un modèle dit « déductif-statistique » où l'explanandum est « déduit » de l'explanans avec une probabilité suffisante pour soutenir l'explication, tandis que dans le modèle « inductif-statistique » la relation entre explanans et explanandum est inductive.

Enfin, il arrive que les conditions du modèle déductivo-nomologique soient insuffisantes à l'explication. Les contre-exemples d'asymétrie explicative (par ex. les lois de l'optique géométrique et l'angle du soleil n'expliqueraient pas la hauteur du poteau) et d'impertinence explicative (par ex. le fait qu'un homme ne tombe pas enceint ne s'expliquerait pas par la régularité voulant qu'un homme prenant des pilules contraceptives ne tombe pas enceint) sont notoires. Dans les deux cas, la connexion entre explication et prédiction est rompue. Une raison, souvent invoquée, pour laquelle de telles prédictions ne seraient pas des explications est qu'une explication exigerait d'identifier une cause.

Mais on objecte aussi à ce modèle de ne pas présenter les conditions nécessaires à l'explication. Ainsi, des énoncés causaux singuliers, comme « l'impact de mon genou sur la table a causé le renversement de mon encrion », permettrait d'expliquer, en sorte que la stricte subsumption sous des énoncés nomiques ne serait pas nécessaire. L'objection notoire à cet argument est celui de la « structure cachée », comme quoi des énoncés causaux singuliers, censés fournir une explication que l'on pourrait qualifier d'ordinaire parce qu'intuitive, ne fournissent toutefois pas une explication complète, idéale ou « scientifique » et se réfèrent, *in fine*, implicitement à un schème déductif répondant au modèle déductivo-nomologique.

Devant les insuffisances du modèle déductivo-nomologique on a donc proposé le *modèle causal-mécanique*, où l'explanans fait appel à des causes ou à des mécanismes (voir Salmon 1984, 1989 ; Bunge 1997 ; Machamer & Darden 2000 ; Woodward 2003b). En général, une telle théorie de l'explication présente une dépendance entre l'explanans et l'explanandum reposant sur une relation d'ordre empirique comme une continuité spatio-temporelle, un échange de quantité physique comme l'énergie. Ce type de théorie est très intuitif, l'état actuel d'un système quelconque étant généralement expliqué par l'évolution temporelle d'un état initial dans une séquence de relations et de relata causaux, ce qui permet d'anticiper, de prédire, de comprendre l'occurrence de l'état final, bref de l'effet.

Pour autant, on ne juge pas les lois comme impertinentes à l'explication. Seulement, des lois strictes et universelles seraient difficiles à découvrir et leur adéquation empirique, surtout dans le cas de lois très générales et très abstraites, serait généralement le résultat de plusieurs correctifs et d'addition d'éléments *ad hoc* (Cartwright 1983). Les lois permettant d'expliquer n'auraient pas non plus besoin

d'être aussi strictes que l'exige le modèle déductivo-nomologique mais seulement, dans une acception plus relâchée, de présenter une certaine dépendance contrefactuelle (discuté plus loin ; Woodward 2003a). Les lois joueraient aussi un rôle important dans l'identification de mécanismes pouvant jouer le rôle d'explanans. Pour Bunge (1997 et 2003), un mécanisme est un ensemble de processus prenant place dans un système concret et produisant ou empêchant des changements. Selon lui, un mécanisme peut être causal, probabiliste, ou les deux, mais il doit être nomique, c'est-à-dire qu'il doit répondre à certaines lois, plutôt que tenir du miracle. Des lois peuvent aussi intervenir pour discriminer les vrais processus causaux, qui peuvent transmettre des influences causales appelées des « marques », des pseudo-processus (Salmon 1984 et 1989). Autrement dit, avec le modèle causal-mécanique, les lois demeuraient nécessaires à l'explication mais leur rôle serait différent que celui du modèle déductivo-nomologique, en ce sens où il encadre le choix de processus causal ou de mécanismes pertinents pour l'explication.

La théorie de l'explication proposée par Woodward (2003a) présente une conception de l'explication causale en termes de manipulations et d'interventions. Elle mérite qu'on s'y attarde puisqu'elle est à l'heure actuelle très influente⁷, même dans le débat sur le concept d'émergence (discuté plus loin). Cette conception permettrait de comprendre l'articulation explicative en fournissant un modèle de dépendance contrefactuelle entre l'explanans et l'explanandum associée à une relation de manipulation et de contrôle. La cause expliquerait l'effet : pourquoi [*effet*] ? Parce que [*cause*]. Cette dépendance contrefactuelle est à la base des « questions-w » : que se serait-il passé si les choses avaient été autrement ? Ainsi, en intervenant, de façon réelle ou hypothétique, l'on serait en mesure de déterminer une relation causale. Dans cette théorie, les relatas causaux ne sont pas des propriétés mais des *variables*. Soit deux variables X et Y , une intervention sur X résulte en un changement de X et de Y . L'intervention ne doit pas être corrélée avec d'autres causes de Y et elle ne doit pas affecter Y via un *chemin* (« *path* ») ne passant pas par X . Un chemin est une entité indéfinie exprimant un lien causal entre deux variables qui sont des propriétés pouvant prendre plus d'une valeur. X est une *cause* de Y si et seulement si la valeur de Y changerait si une intervention sur X avait lieu dans certaines circonstances V . De sorte qu'une intervention I fixe la valeur de la variable X qui produit un changement de la variable Y ou de la distribution de probabilité de Y tandis que les autres variables Z_i de V restent fixées. Formellement, on a

$$I \rightarrow X \rightarrow Y.$$

⁷ Cette théorie causale n'est pas entièrement nouvelle car elle rejoint le principe scholastique *causa cessante cessat effectus*, mais aussi la formulation de la causalité par Galilée (voir Bunge 1959 : 33).

Pour qu'une généralisation exprime une relation causale, la description doit rester vraie, c'est-à-dire invariante sous certains types de changements. Cette condition d'invariance se traduit plus largement à l'ensemble des représentations causales car elles doivent satisfaire la condition de *modularité* où, en termes simples, l'intervention sur une variable d'une représentation, comme une équation mathématique, ne modifie pas les autres représentations causales.

Ainsi, la stratégie de la théorie manipulationniste de la causalité de Woodward (2003) se décline succinctement comme suit :

STRATÉGIE M. Le changement de valeur d'une variable T ou de la distribution de probabilité de T (explanandum) dans un contexte V est expliqué si et seulement s'il y a une intervention fixant la valeur de X produisant le changement exogène de T lorsque les autres variables de V sont fixées, tel qu'un chemin direct et unique (modulaire), et orienté vers le futur, passe de X à T (explanans).

Cette théorie a un caractère intuitif et semble ménager les suspicions humiennes. L'intervention semble apporter une « réalité » aux relata causaux, en supposant que l'intervention soit elle-même réelle, du moins l'inférence quant à leur réalité reçoit un appui phénoménologique. Le caractère explicatif des inférences causales repose essentiellement sur la dépendance contrefactuelle d'une manipulation d'une variable antérieure (cause) au changement d'une autre variable (effet). Bien que cette dépendance contrefactuelle s'éloigne des desiderata humiens, elle répond à ce desideratum crucial qui distingue les lois des non-lois dans le modèle déductivo-nomologique ; mais ici elle renvoie plutôt aux généralisations invariantes.

La littérature philosophique sur ce qu'est une cause foisonne (par ex. Bunge 1959 ; Mackie 1974 ; Schaffer 2006 ; Jodoin 2010a). Plusieurs ont contesté que les sciences et au premier chef la physique avaient besoin de la causalité en général (Russell 1913) ou qu'elle cherchait à satisfaire les desiderata d'une explication causale (Thalos 2002). Une objection bien connue au modèle causal-mécanique est la difficulté d'identifier clairement ce qui est causalement pertinent, parmi la vaste gamme de facteurs impliqués dans la description d'un phénomène, à l'explication. Encore une fois, les lois scientifiques peuvent encadrer jusqu'à un certain point ce choix.

Les critiques des théories orthodoxes – déductivo-nomologique et causal-mécanique – ont surtout porté non pas sur la *pertinence* de l'approche, mais plutôt sur sa *nécessité* étant donné les exigences qu'elles placent sur l'explication. C'est-à-dire qu'un explanans exprimé selon les termes de ces théories, comme une loi ou une cause, serait suffisant, dans des circonstances particulières, mais pas nécessaire en toutes circonstances à une « bonne » explication, une explication satisfaisante. C'est en ce sens qu'elles définissent une orthodoxie de l'explication scientifique, d'une part, mais aussi, d'autre part, des cas

particuliers d'un ensemble plus vaste de stratégies explicatives. C'est donc pourquoi des stratégies explicatives alternatives seraient non seulement envisageables mais aussi souhaitables en regard de certains objectifs louables (prochaines sections).

2.1.1.1 EXPLICATION PROBABILISTE

Le portrait global d'une telle stratégie ressemble à ceci (Sklar 1973 et 1993) : soit une généralisation statistique S , si les conditions C_i qu'elle décrit sont remplies, alors, avec une probabilité p , l'événement décrit par E se produira. Cette généralisation peut être déductive ou inductive. Il y a plusieurs desiderata (qui ne sont pas tous compatibles) pour cette stratégie : p doit être élevée (généralement supérieure à 0,5) ; l'explanans doit augmenter p ; p doit être générée de S à partir d'une classe de référence ; p produit un « degré de croyance » (suffisamment) justifié. Il n'y a pas de consensus sur une théorie de l'explication statistique qui permettrait de réconcilier tous ces desiderata. Mais, il est généralement admis (quoique contestable, bien sûr) que l'« expectative » de E à partir de la combinaison $\{S, C_i, p\}$ possède un pouvoir explicatif. Par exemple, à la question « pourquoi telle occurrence ? », une réponse satisfaisante pourrait être « parce sa probabilité était élevée ». L'effectivité de cette expectative repose en grande partie sur l'interprétation des probabilités que l'on adopte (Chapitre 5).

Il est aussi possible, à l'inverse, d'aller plus loin en inférant de cette expectative la « cause » de E , comme c'est le cas de certaines stratégies causales probabilistes. Par exemple, certains auteurs (Hamilton 2009 : 5) affirment que si l'observation révèle que l'expectative était fondée, alors le schéma explicatif recèle la « cause » du phénomène observé. En mécanique statistique, où les probabilités sont appliquées à des systèmes, les molécules, que l'on suppose *in fine* déterministes, on retrouve essentiellement deux types de stratégie explicative, l'une pour les systèmes à l'équilibre et l'autre pour les systèmes qui ne sont pas à l'équilibre (Sklar 1993). Dans ce dernier cas, la stratégie est soit causale, où une force extérieure dévie le système de l'équilibre, soit déductivo-nomologique, où il y a une tendance vers l'équilibre qui est stipulée. Dans le premier cas, la stratégie est moins intuitive : l'équilibre, qui est tenu pour acquis et qui est supposé être lié à certains états théoriques, doit, en présupposant certaines hypothèses plausibles, avoir telles et telles caractéristiques particulières.

2.1.1.2 FICTIONS ET IDÉALISATIONS

Certaines théories scientifiques, pour ne pas dire la majorité, ont recours à des *fiction*s particulières ou à des *idéalisations*. Les idéalisations abondent en sciences : gaz parfaits, systèmes isolés, agents (parfaitement) rationnels, stratégies évolutionnaires stables, etc. Mais leur rôle demeure largement incompris. L'un des problèmes qu'elles posent est qu'elles semblent rendre inopérant le modèle

déductivo-nomologique selon lequel une théorie est un schème déductif utilisé avec des lois et des conditions initiales afin de dériver des prédictions ou des explications. Or un schème déductif devrait rendre des conclusions vraies lorsque les prémisses le sont aussi. Mais si les idéalizations sont admises comme prémisses et qu'elles sont par définition fausses, comment une théorie peut-elle rendre des conclusions vraies et donc prédire et expliquer ?

Ainsi, les sciences opèrent parfois sur le mode des *comme si*. Pour autant, le discours scientifique n'est pas pur mensonge. Contrairement aux fictions, qui sont aussi fréquemment utilisées en sciences, les idéalizations approximeraient les cas réels et possèderaient donc un caractère de « vérisimilitude » (Popper 1963), en ce sens où elles sont constitutives de certaines propositions susceptibles d'être corroborées ou réfutées. Il peut en effet y avoir des arguments permettant d'évaluer que nous ayons fait un progrès vers la vérité⁸ (Niiniluoto 2007). Bunge distingue ainsi les « fictions fantaisistes » (« *wild fictions* »), comme les divinités ou les licornes, des « fictions maîtrisées » (« *tame fictions* »), comme le concept de fonction ou celui de particule simplifiée. Ainsi, contrairement à un cas purement fictif, un cas idéal (fictions maîtrisées) est analogue au cas réel et il peut être approché graduellement (Weisberg 2007).

La distinction entre idéalisation et fiction (fantaisiste) peut alors être considérée comme un continuum où une simple fiction, si elle possède certaines caractéristiques, peut être une idéalisation. En ce sens, le caractère fictif se présenterait par degrés (Russell 1948). Ce que l'on qualifie de « fiction utile » aurait ainsi un certain degré de réalité et de vérité. Bien que le sens précis de cette expression soit difficile à établir, on peut affirmer qu'une fiction présentant les caractéristiques adéquates permet des prédictions avérées ou, plus précisément, ses anticipations ne seraient pas démenties. Après tout, une fiction qui expulserait toute trace de l'univers non fictionnel, autrement dit tout élément se rapportant à la réalité, deviendrait simplement incompréhensible. Ce qui caractérise une idéalisation ou un « fiction maîtrisée » n'est donc pas de nous livrer une description fidèle de la réalité mais plutôt de nous permettre de nous projeter dans la situation qu'elle représente, de livrer un gain cognitif intéressant. Entre autres, elle peut offrir une description plus simple, faisant l'économie de mots par rapport à une description littérale.

Cette idée s'appuie en définitive sur quelque chose comme une « logique de l'approximation » (Ben-Menahem 2000). L'idée est de prolonger la logique binaire à une logique à valeur réelle, avec le concept de « vérité partielle », où les valeurs de vérité peuvent se situer entre ces extrêmes de la vérité

⁸ En supposant que la vérité n'est pas directement reconnaissable (la connaissance est alors inférée d'une croyance et d'une justification), le défi pour la connaissance consiste à trouver de bonnes justifications, et ce progrès est alors évalué à l'aune de différents ensembles (contrastifs) de justifications (Jodoin 2014a, section 2.2.1).

et de la fausseté absolues (Bunge 2003). Par exemple, l'énoncé « π égale 3 » est plus vrai que « π égale 2 ». De sorte que si les prémisses sont suffisamment vraies, alors la conclusion est très près d'être vraie. Similairement, les idéalizations, qui ne peuvent prétendre par définition à une description « absolument vraie » de la réalité, seraient « suffisamment vraies » pour qu'elles puissent faire partie d'un schème déductif valide selon cette logique modifiée. L'idée de la logique de l'approximation recouvre l'intuition que des descriptions plus « fines », plus détaillées, sont plus vraies que des descriptions plus « grossières », contenant moins de détails. Par exemple, il semble rationnel de dire, à propos de ma position géographique, que l'énoncé « je suis à Paris » est plus vrai que l'énoncé « je suis dans l'hémisphère nord » ; ou encore, d'un homme mesurant deux mètres, que l'énoncé « cet homme mesure entre 1,90 et 2,10 mètres » est plus vrai que l'énoncé « cet homme mesure plus d'un mètre ». Des précisions s'imposent, toutefois. D'abord, à prime abord, les concepts d'approximation et de vérité partielle ne semblent pas commis à une théorie particulière de la vérité (et, de plus, une définition de la vérité ne fournit pas nécessairement les moyens de l'identifier). Ensuite, le niveau de détails n'est pas suffisant pour garantir la vérité de l'énoncé descriptif : une somme d'énoncés faux ne fait pas un énoncé complexe vrai.

Les exemples de « cas limite » ou de « cas asymptotique » illustrent bien ce caractère approximatif des idéalizations. En général, un paramètre théorique représentant une propriété d'un système est susceptible de varier dans certaines conditions de sorte que, lorsqu'il se rapproche d'une certaine valeur limite ou asymptotique, certaines représentations des propriétés du système ou la description de son comportement global adoptent des valeurs ou une tendance particulières. Et en général, la valeur asymptotique de ces paramètres est soit zéro, soit l'infini. Un exemple patent est celui de la « limite thermodynamique » lorsque le nombre de molécules d'un système tend vers l'infini (Chapitre 4). Une particularité intéressante de cette limite est qu'elle semble nécessaire pour simuler ou prédire les phénomènes de transition de phase puisque le traitement mathématique de systèmes mécaniques finis ne rend pas compte de ce genre de phénomènes. Encore plus intéressant : ce serait une raison pour qualifier ces phénomènes d'« émergents » (Liu 1999 ; Batterman 2000 ; plus de détails à la section 2.3).

L'objectif des idéalizations de type asymptotique est parfois simplement méthodologique en ce qu'elles permettent de simplifier les calculs ou simplement de les réaliser. Par exemple, un processus mécanique peut être considéré sans friction, ce qui est pratiquement impossible, de sorte que les effets dissipatifs dus à cette friction peuvent être négligés. Aussi, ce qu'on appelle les « effets relativistes », responsables de la dilatation du temps et de la contraction de l'espace, exigeant ainsi des corrections cinématiques, sont perceptibles, mesurables, qu'à très grande vitesse, de sorte qu'à plus faible vitesse ils peuvent être négligés. Il arrive aussi que la mise en application de certaines lois exige de transformer jusqu'à un certain point les propriétés des systèmes auxquels elles sont censées s'appliquer pour être en

mesure de les mettre en équation. C'est le cas notamment d'une planète considérée comme un point matériel.

Mais leur rôle devient particulièrement intéressant au sein d'explication dont la stratégie consiste pour l'essentiel à montrer en quoi des événements différents peuvent mener au même phénomène, que certaines modifications dans la description de l'état initial d'un système conduit à la même prédiction quant à l'état final, bref qu'il y a une tendance exemplifiée par l'idéalisation. Par exemple, il y a une tendance, décrite par la première loi de la mécanique newtonienne, des corps à se mouvoir selon une trajectoire rectiligne uniforme, un phénomène qui, strictement, n'est jamais exemplifié ; lorsque ce n'est pas effectivement le cas, pour ainsi dire *toujours*, on fait appel à des causes, des forces, pour expliquer l'écart qu'il peut y avoir avec cette tendance vers ce cas limite. Or, plusieurs ensembles de forces peuvent mener au même écart d'avec une trajectoire rectiligne uniforme (par exemple, à la même accélération), tout comme plusieurs éliminations successives de certaines forces appliquées à divers corps tendent vers un mouvement rectiligne uniforme. Dans le cas de la loi des grands nombres dans sa version faible, lorsque le nombre d'événements associés à une variable aléatoire tend vers l'infini, il devient de moins en moins probable, jusqu'à la limite impossible puisque la probabilité devient nulle, que la valeur de cette variable diffère de l'espérance mathématique. Dans certains cas plus subtils, l'idéalisation asymptotique mène à ce qu'on appelle la réalisabilité multiple (section 2.2.5). Ce type d'idéalisation peut servir alors à expliquer pourquoi certaines propriétés sont instanciées de plusieurs façons, pourquoi elles sont « universelles » (Batterman 2000).

Les fictions peuvent donc se présenter sous diverses formes et avec diverses vertus épistémologiques. Elles peuvent passer de fiction fantaisiste à fiction maîtrisée, ou utile, autrement dit à des idéalizations jouant un rôle dans certaines stratégies explicatives. Leur statut simplement logique n'est donc pas ce qui les caractérise, parce qu'elles sont en un sens strict fausses, mais elles se rapprochent de propositions pouvant être vraies, s'apparentant à des hypothèses qui ne sont pas réfutées sans être avérées. Elles permettent aussi de comprendre certains phénomènes en nous projetant dans des situations pourtant irréelles. Puisqu'elles se présentent souvent comme des représentations auxquelles des éléments caractéristiques auraient été omis, autrement dit des simplifications, des abstractions, elles contribuent à former des énoncés abstraits certes, mais généraux. Or, si la « généralité et la simplicité sont la substance de l'explication » (Cartwright 1983 : 112), il ne faut s'étonner de retrouver des idéalizations en sciences. Pour Bunge (2006 : 10), les idéalizations (« fictions maîtrisées ») constitueraient « le chemin vers et à partir de la réalité »⁹. Il n'y a de science qu'au général disait déjà Aristote.

⁹ « *the path to and from the reality* ».

2.1.2 Alternatives

Tel que mentionné, l'orthodoxie de la théorisation de l'explication scientifique qui s'est formée au milieu du XX^e siècle fait généralement consensus quant à l'identification de modèles d'explication pertinents et efficaces, mais plusieurs auteurs lui reprochent d'omettre certains types d'explication utilisés en sciences et qui sont supposément valides. Ainsi, on veut bien admettre qu'une loi ou une cause soit en mesure d'expliquer, supposant que ces concepts soit bien définis, mais il semble raisonnable de dire qu'il est possible d'expliquer sans avoir recours à des lois ou à des causes. Tant et aussi longtemps que le rôle d'une explication ne sera pas clairement établi, il faut s'attendre à rencontrer une pléthore de modèles d'explication.

Le cas de la biologie est particulièrement intéressant puisqu'elle est considérée comme une « science spéciale », c'est-à-dire, croit-on, comme une science dont les entités qu'elle décrit sont composés d'entités physiques, généralement jugées plus fondamentales pour une raison ou pour une autre. Cette appellation se rattache donc à une conception réductionniste. Celle-ci n'implique toutefois pas que les outils épistémologiques pour décrire et expliquer le comportement de ces entités proprement biologiques se réduisent effectivement à ceux de la physique.

Dans le cas de la biologie, donc, Ernst Mayr (Sober 1993 : 6) a proposé une distinction entre *explication proximale* (*proximate explanation*) et *explication ultime* (*ultimate explanation*). Dans le premier cas, l'explication est dynamique et porte davantage sur les causes immédiates et les mécanismes sous-jacents au phénomène à expliquer, généralement au niveau individuel ; elle concerne des disciplines comme la physiologie et l'ontogénie. Dans le second cas, l'explication est plutôt statistique et cherche ses assises dans le passé lointain (d'une espèce par exemple) et dans le récit historique et évolutionnaire de l'entité en question, souvent au niveau populationnel ; elle est susceptible d'intéresser principalement la théorie de l'évolution et la phylogénétique. On constate donc que la distinction résume, en partie seulement, la différence entre les approches explicatives faisant appel à des mécanismes causaux et celles consistant en une subsomption sous des énoncés généraux, vus précédemment. En effet, une explication proximale réfère typiquement à des causes bien précises, comme la pression osmotique dans une cellule par exemple, voire même à ce qui peut sembler contingent, par exemple dans un cas de dérive évolutionnaire avec la mort accidentelle d'une gazelle autrement adaptée à son milieu. Tandis qu'une explication ultime ne se réfère pas aux séquences particulières d'événements, aux « chemins étiologiques », souvent perdus dans l'écume des jours, mais offre plutôt un cadre plausible aux explications proximales en faisant généralement appel à des mécanismes adaptationnistes, à la sélection naturelle.

Mais si la biologie, ou plutôt les sciences biologiques, font appel à ces deux stratégies explicatives, est-il raisonnable de penser que leurs explications ont quelque chose de particulier, de différent de l'orthodoxie de l'explication ? Selon Sober (1984 : 147), avec Darwin se présente une « innovation explicative » qui implique un changement logique et non pas un changement conceptuel. L'innovation de Darwin est d'avoir situé l'explication au niveau de la population où l'explanans ne repose pas sur le changement individuel (*explication développementale*) mais sur un statisme individuel combinée à une sélection individuelle (*explication variationnelle*). Il y a ainsi un nouveau « contexte contrastif » entre, d'une part, les explications présentant la séquence des étapes par lesquelles est passé un organisme et celles, d'autre part, se concentrant sur un ensemble d'individus différents, c'est-à-dire présentant de petites variations entre eux, lesquelles sont hérissables, et soumis à une pression de sélection. Or, ces deux approches exigent des hypothèses *historiques*. Pour Gould (1986 : 60) aussi l'innovation explicative du darwinisme se situe au niveau de la population et, plus précisément, elle montre le « triomphe de l'homologie sur les autres causes d'ordre » ; un homologue étant le résultat d'une « histoire phylogénétique continue » (Ereshefsky 2012 : 384). L'idée n'est pas tant que l'explication biologique est inopérante sans l'évolution, mais plutôt d'insister sur son incomplétude sans l'évolution. Et puisque l'évolution au sens darwinien implique des hypothèses historiques, l'explication biologique anhistorique est incomplète. Comme le souligne Falk (2009 : 3), les affirmations de régularité et de nomicité des sciences biologiques sont subordonnées aux événements passés qui ont été déterminés, partiellement du moins, par la sélection naturelle et les contraintes de structure et de fonction alors présentes¹⁰.

Rosenberg (2001) a repris cette idée du caractère historique de l'explication biologique. Selon lui, puisque la biologie présente des lois fonctionnelles (par ex. tous les amphibiens se reproduisent sexuellement) et puisque la sélection naturelle est « aveugle » à la structure physique, une loi biologique fonctionnelle présentera des entités physiquement diverses, en conséquence de quoi la biologie n'aurait pas de lois strictes. Mais la biologie peut éviter les obstacles à la découverte de lois ainsi que les problèmes légués par la conjonction de l'individuation fonctionnelle et de la sélection naturelle, si elle abandonne l'individuation fonctionnelle et adopte l'identification structurelle. Le prix est cependant trop cher payé : les lois structurelles seraient à toutes fins pratiques inutiles aux prédictions à notre échelle temporelle ou aux explications intelligibles et d'intérêt. En effet, abandonner l'individuation fonctionnelle revient à troquer la biologie pour la chimie organique, autrement dit à perdre son caractère proprement biologique. Cela n'est pas dramatique pour une entreprise de réduction de la biologie à la chimie ou à la physique. Mais, si l'explication biologique est narrative comme celle en

¹⁰ Les entreprises explicatives, que l'on retrouve surtout en génétique quantitative et en génétique des populations, et se basant sur l'hypothèse que ces systèmes peuvent être modélisés comme des processus markoviens doivent donc faire face à cette contrainte, à cette objection de l'historicité des phénomènes biologiques.

histoire, et que l'explication narrative n'a pas besoin de lois, alors le problème de l'explication biologique est résolu. L'explication biologique, comme celle historique, propose non pas une déduction nomologique mais plutôt une série intelligible d'occurrences dans laquelle prend place l'événement à expliquer, l'explanandum. Or, les principes fondamentaux de la théorie darwinienne sont universels et nomologiquement nécessaires et constituent les seules lois (*ceteris paribus*) de la biologie. Ce serait l'application de ces lois aux conditions initiales des systèmes d'intérêt en biologie qui engendre les « types fonctionnels », lesquels font en sorte que le reste de la biologie est implicitement historique.

L'identification du caractère historique de l'explication biologique n'en donne évidemment pas une caractérisation complète. Sans atteindre l'exhaustivité, l'innovation explicative dont parle Sober se concentre sur la sélection naturelle darwinienne et comporte deux étapes essentielles (Mayr 1997) : la première est la production d'une grande quantité de variations servant de matériel à la seconde, soit le processus de sélection ou d'élimination. Or, d'une part, ce processus de sélection est généralement le résultat d'interactions complexes entre l'organisme et son environnement, lequel change de manière contingence ou aléatoire (du moins par rapport à la théorie). Mais les régularités contingentes peuvent néanmoins avoir un pouvoir explicatif selon Brandon (1996). D'autre part, cette interaction est reflétée par une propriété relationnelle bien connue mais pourtant mal comprise, la fitness (ou valeur sélective). Cette notion est fondamentale puisqu'elle est au centre de la stratégie darwiniste expliquant la diversité, la complexité et l'adaptation du domaine biologique (Bouchard & Rosenberg 2004). Bien que la fitness soit généralement définie comme une caractéristique d'un phénotype individuel, en pratique elle est uniquement mesurée et appliquée aux groupes d'individus. Ainsi, selon la terminologie de la génétique quantitative, la fitness est la valeur génotypique des fitness individuelles d'un groupe ayant en commun un même génotype. La fitness moyenne (W_m) est donc une mesure de la contribution collective moyenne (le nombre de gamètes transmis) à la prochaine génération et non la contribution particulière d'un certain individu. Cette transposition de l'individu au groupe peut être justifiée en théorie de l'évolution si l'évolution est définie en termes de changements génétiques d'une génération à l'autre ; par exemple, les effets moyens mesurent les variations phénotypiques transmises d'une génération à l'autre et sont fonction uniquement des valeurs génotypiques du groupe, et non de l'individu. Ce n'est pas tant que les phénotypes individuels ne sont pas importants, mais plutôt qu'ils n'apportent rien de plus, de ce point de vue, à l'explication des changements évolutionnaires. Dans le cas d'un génotype qui est établi (c'est-à-dire mesuré), la représentation de l'action de la sélection naturelle est déterminée (agit-elle ou non, et dans quel sens ?) par celle de l'action de l'excès moyen (un paramètre typiquement populationnel), c'est-à-dire par la déviation génotypique moyenne qu'un type spécifique de gamète est censé présenter. De ce point de vue, contesté toutefois, l'expression « la survie du plus apte » (« *survival*

of the fittest ») est fausse si elle réfère aux individus (organismes) et à leur phénotype, car la sélection naturelle « favorise » les gamètes avec un excès moyen positif pour la fitness (Templeton 2006 : 349).

Ce caractère historique de l'explication biologique est ainsi largement conditionné par le degré de précision descriptive de l'explanandum, aux descriptions à « grenage fin » *versus* celles à « grenage grossier ». Un explanandum très précis, très spatio-temporellement localisé, bref à « grenage fin », exige généralement un explanans contingent, une séquence d'événements bien particuliers, tandis qu'un explanandum plus « large », décrit en termes plus « grossiers », appelle un explanans qui semble inclure une certaine nécessité en ce qu'il montre que ce à quoi réfère l'explanandum devait se produire ou à tout le moins se produit dans plusieurs circonstances différentes (Huneman 2010). Un exemple d'un explanandum à « grenage fin » serait la disparition d'une espèce de poisson dans un lac au Michigan, tandis qu'un exemple d'un explanandum à « grenage grossier » serait la diversité des espèces sur terre.

Il faut alors noter deux choses. D'une part, la situation n'est pas unique à la biologie et elle se présente aussi dans d'autres disciplines comme la physique. Par exemple, l'explication de *cette* vitesse de *tel* projectile exige, en plus des lois mécaniques appropriées, les conditions initiales de ce système et possiblement tout un réseau d'événements contingents comme, par exemple, le rebond sur telle surface ayant telle élasticité, etc. D'autre part, l'adéquation de l'explanans avec l'explanandum en fonction du degré de précision descriptive est asymétrique, en ce sens où un explanans à « grenage grossier » peut bien satisfaire un explanandum à « grenage fin » mais un explanans à « grenage fin » ne peut généralement pas satisfaire un explanandum à « grenage grossier ». On peut, par exemple, invoquer la sélection naturelle, un explanans à « grenage grossier », pour expliquer la disparition d'une espèce de poisson dans un lac au Michigan, mais il ne fait aucun sens de parler de l'action d'un certain prédateur dans ce lac, soit un explanans à « grenage fin », pour expliquer la diversité des espèces. Cependant, même si l'adéquation entre explanandum et explanans opère, c'est-à-dire qu'il soit vrai que la sélection naturelle s'applique à cette situation particulière du lac au Michigan, cet explanans et par conséquent cette explication risquent d'être largement insatisfaisants. Le degré de précision descriptive doit donc être adapté à celui de l'explanandum.

L'explication de type historique apparaît importante lorsqu'une incomplétude est manifeste dans le récit reliant l'explanandum à son explanans, c'est-à-dire lorsque plusieurs causes semblent expliquer ou sont à tout le moins compatibles avec l'occurrence d'un même phénomène. Il s'agit alors d'avoir recours, ici aussi, à une série intelligible d'occurrences dans laquelle prend place l'explanandum, donc d'offrir une explication de type historique. Ainsi, dans le cas où un système est dans un état singulier qui

n'est pas déductible de certaines propriétés générales, comme sa composition chimique, et des conditions aux limites, et que d'autres états lui étaient également accessibles,

la seule explication est donc historique, ou génétique : il faut décrire le chemin qui constitue le passé du système, énumérer les bifurcations traversées et la succession des fluctuations qui ont décidé de l'histoire réelle parmi toutes les histoires possibles. » (Stengers & Prigogine 1979 : 231).

Une objection à cette stratégie serait de l'accuser d'être simplement descriptive et non pas explicative. Le repli consiste généralement à décrire le facteur causal le plus pertinent ou le plus influent (Humphreys 1989). Une stratégie largement employée est d'identifier la cause ou la séquence d'états la plus *probable*.

La probabilité apparaît ici comme la conséquence de la détermination d'une gamme de possibles pour un phénomène particulier, même si le choix d'une mesure de probabilité, nécessaire à l'attribution d'une valeur, puisse être difficilement justifié (Chapitre 5). Plus particulièrement, l'identification d'un sous-groupe de possibles « plus probables » parmi cette gamme de possibles permet de conférer une « expectabilité » (discuté plus loin) à ce sous-groupe. Et l'expectabilité – comme quoi l'on est en droit de s'attendre que ce qui s'est passé devait se passer – permet généralement de définir une explication. À l'instar de l'objection que l'on adresse au modèle déductivo-nomologique et à l'obligation d'avoir une loi, l'expectabilité ne semble pas nécessaire à l'explication même si elle paraît suffisante. L'effectivité de cette stratégie explicative réside dans une interprétation satisfaisante de la probabilité, de cette clause « plus probables », afin de soutenir cette expectabilité, ce qui est le cas notamment avec les interprétations propensionniste et fréquentiste.

On constate toutefois que cette stratégie explicative est compatible avec une approche causaliste. Offrir une série intelligible d'occurrences dans laquelle prend place l'explanandum peut ainsi se résumer à présenter une séquence causale. Qu'il y ait un ensemble de séquences possibles selon les informations disponibles et circonscrivant l'explanandum, et que certains critères, comme la pertinence ou la forte probabilité, visant à sélectionner l'une d'entre elles, ne compromettent pas radicalement l'approche causaliste. Sober (1983) contraste les explications faisant justement appel à des séquences réelles d'événements menant à un certain résultat, qu'il qualifie (justement) de causales, des explications où le résultat est expliqué en montrant qu'un très grand nombre de conditions initiales d'un système évolueraient vers cet état résultant que l'on souhaite expliquer, qu'il appelle des *explications d'équilibre* (« *equilibrium explanations* »). Il prend pour exemple Fisher (1931) qui montre que si une population dévie d'une répartition sexuelle 1 : 1, il y aura un avantage reproductif favorisant les parents surproduisant le sexe minoritaire, de sorte que cette répartition 1 : 1 sera le point d'équilibre résultant. Ce type d'explication est ainsi typiquement employé en génétique des populations où un paramètre comme la

fréquence d'allèles aboutit à un point d'équilibre sous l'effet combiné de plusieurs forces et selon plusieurs combinaisons de ces forces (Chapitre 6).

Cette stratégie explicative est très semblable à ce que propose Batterman (2000 et 2002) pour expliquer l'« *universalité* », c'est-à-dire un comportement similaire chez différents systèmes (Weslake (2010) affirme d'ailleurs qu'il s'agit d'un cas particulier d'une explication d'équilibre). Il soutient d'une part que la *réalisabilité multiple* (section 2.2.5) est un cas particulier de l'universalité et, d'autre part, qu'il existe une opération mathématique, soit un groupe de renormalisation appliqué à certains systèmes physiques, montrant que les détails de la microstructure de ces systèmes est largement impertinente (c'est-à-dire qu'ils ne font pas de différence) à leur comportement à peu près identique. Cette stratégie explicative est donc la même, conceptuellement, que celle des explications ultimes (vues précédemment) où plusieurs séquences causales sont possibles entre un certain état initial, une structure biologique naissante par exemple, et un état final correspondant à un stade plus avancé ou plus évolué de cette même structure. Pour Rosenberg et McShea (2008 : 117), il s'agit d'un cas de réalisabilité multiple, c'est-à-dire que plusieurs structures peuvent exemplifier les mêmes mécanismes d'adaptation.

Il y a évidemment d'autres stratégies explicatives bien différentes des stratégies nomologiques, causales ou encore de type historique (causaliste ou probabiliste). Ainsi des stratégies explicatives en sciences font appel à des éléments abstraits, mathématiques, ou encore, comme on l'a vu, à des idéalizations. Ainsi, des symétries peuvent jouer un rôle explicatif en imposant des contraintes (de symétrie justement) aux lois physiques et ainsi expliquer la forme de ces lois ou l'occurrence de certains phénomènes (Brading & Castellani 2003). Et puisqu'une unification est parfois considérée comme une exigence d'une bonne explication, le pouvoir unificateur des symétries peut aussi être perçu comme explicatif.

Ainsi le *modèle unificationniste*, d'abord proposé par Friedman (1974) puis défendu par Kitcher (1981), tente de montrer en quoi l'explication d'un phénomène est reliée à sa compréhension. L'idée générale est de fournir un portrait du monde simplifié via une réduction du nombre de phénomènes que l'on considère comme « ultimes », comme des « faits bruts ». Kitcher (1989 : n17) cite Huxley pour qui la réussite de Darwin est d'avoir réduit au maximum le nombre d'incompréhensions fondamentales. Une explication consiste ainsi en une unification définissant l'ensemble des arguments où est minimisé le nombre de prémisses utilisées et où est maximisé le nombre de conclusions obtenues. Ce modèle est intuitif car une théorie qui unifie plusieurs phénomènes est, en un sens assez évident, plus générale et sans doute avec un « pouvoir systématique » plus grand (prochaine section). Il peut aussi s'appuyer sur des considérations historiques comme l'unification des lois célestes et terrestres avec la physique newtonienne ou encore celle des phénomènes électriques et magnétiques dans l'électromagnétisme

maxwellien. Cependant, non seulement ce modèle ne justifie pas la thèse à l'effet que notre compréhension du monde soit correctement mesurée par le degré d'unification théorique, mais en plus il soutient implicitement que l'explication découle de la compréhension d'un explanandum, entendu comme phénomène indépendant, subsumé sous un explanans, soit un ultime point d'appui comme « fait brut », qui lui demeure complètement mystérieux, incompris (Barnes 1992). Ce modèle a aussi maille à partir avec les asymétries explicatives, avec le fait que certaines unifications statistiques ou mathématiques ne semblent pas expliquer, et avec cette conséquence (« *winner-take-all conception* ») que seule la théorie la plus unificatrice puisse être explicative (Woodward 2003a).

Devant cette pléthore de modèles, de théories de l'explication et des objections qu'on leur adresse, il n'est pas étonnant que l'on ait argué pour un pluralisme. Ce pluralisme a pris la forme d'un *pragmatisme* chez van Fraassen (1980) qui soutient qu'une explication est une relation non pas entre explanandum et explanans mais entre question et réponse, lesquelles sont contextualisées. Pour qu'une certaine proposition puisse être une réponse à et donc expliquer une autre proposition, elle doit être plus pertinente que d'autres réponses et être compatible avec un contexte de connaissances scientifiques. Sa position n'est pas celle de l'instrumentalisme, qui nie l'existence des entités inobservables, mais plutôt celle de l'*empirisme constructivisme*, qui suspend son jugement sur ces entités et demeure donc agnostique à leur endroit. L'objectif de la science serait dans ce cas de construire des modèles jugés *empiriquement adéquat* et non la découverte de la vérité sur les inobservables.

On constate donc la diversité des stratégies explicatives. Cette diversité est en quelque sorte internalisée par l'approche pragmatiste (que je ne discuterai pas plus avant ici). Elle laisse toutefois de côté tout aspect normatif de l'explication, car pour éviter la trivialité où toute proposition ou presque peut être jugée pertinente, elle repose en définitive sur le discours scientifique.

2.1.3 Réduction

Le progrès scientifique est souvent considéré soit comme l'addition de faits nouveaux, soit comme le résultat d'une unification de plus en plus large de descriptions de phénomènes divers par des concepts de plus en plus complets, généraux et, en un certain sens, plus simples. Ce second résultat est généralement obtenu en assimilant, en *réduisant* un appareil théorique, avec ses concepts et ses lois, à un autre appareil théorique jugé, justement, plus complet, plus général et peut-être plus simple. Le concept de réduction peut ainsi s'appliquer soit à l'activité, pouvant être spatio-temporellement localisée, de scientifiques passant d'une théorie ou d'un concept à d'autres jugés supérieurs, soit comme une

relation, épistémologique ou métaphysique, indépendante de ce que peuvent faire les scientifiques¹¹. Dans ce dernier cas, qui fera l'objet de la prochaine discussion, l'idée de base de la réduction est exprimée par le prédicat relationnel « n'est rien de plus que ». Il s'agit alors en quelque sorte de ramener la diversité à l'unité. Il faut donc spécifier et justifier quels sont les *relata*, c'est-à-dire à quoi s'applique la réduction, mais aussi la *relation* elle-même, c'est-à-dire les conditions que les *relata* pertinents doivent satisfaire afin d'instancier cette relation et donc ce prédicat relationnel.

Le réductionnisme est donc une thèse faisant la promotion des réductions. Et comme c'est souvent le cas en philosophie des sciences, elle se présente sous deux formes génériques. D'une part, les *relata* peuvent être des théories, des discours, des modèles ou des concepts, auquel cas la position est épistémologique ou méthodologique. Il s'agit alors d'une position selon laquelle il est désirable, voire même nécessaire, dans l'entreprise scientifique, de procéder à des réductions consistant à absorber, subsumer ou expliquer l'un de ces *relata* par un autre. Sa popularité découle sans doute de la croyance en son prétendu pouvoir à tout expliquer, à tout unifier la diversité des phénomènes et du savoir. Une réduction est en effet similaire à l'analyse en ce qu'elle fait appel aux composantes, plus élémentaires, d'un tout, plus complexe. En ce sens, elle favoriserait la compréhension. D'autre part, le réductionnisme peut être une position ontologique stipulant que la diversité des phénomènes se réduit, se ramène à un seul type d'entités ou un petit nombre d'entités jugées fondamentales. Les *relata* sont, dans ce cas, des entités, des événements, des propriétés (causales). Plusieurs approches réductionnistes sont sympathiques à un certain type de matérialisme ou de physicalisme, comme quoi les objets du monde seraient fondamentalement physiques ou matériels. Cependant, le réductionnisme *per se* est ontologiquement neutre. Le concept de réduction n'est pas commis à une conception ontologique particulière. Mais il est vrai que le réductionnisme retombe souvent sur une conception fondamentaliste comme le physicalisme, stipulant que tous les phénomènes peuvent, *in fine*, s'expliquer par la physique. Par contre, ces deux approches, épistémologique et ontologique, paraissent inexorablement intriquées puisqu'il semble impossible de justifier une assertion réductionniste ontologique sans faire appel à nos connaissances, à nos moyens épistémiques.

2.1.3.1 TAXONOMIE RÉDUCTIONNISTE

Silberstein (2002) présente une taxonomie fort pertinente des divers types de réductionnisme, entre autres parce qu'il la structure comme le « miroir » de l'émergentisme (discuté plus loin). C'est-à-dire que le réductionnisme est généralement en mesure de présenter les conditions nécessaires et suffisantes

¹¹ Van Riel & van Gulick (2014) distinguent ainsi une réduction *diachronique* d'une réduction *synchronique*. L'idée d'une réduction diachronique n'est pas tant qu'il s'agisse de quelque chose de temporel mais plutôt d'une succession, du remplacement d'une théorie par une autre.

à une réduction, qu'elle soit ontologique ou épistémologique, tandis que les cas d'émergence sont souvent présentés comme l'échec de certain type de réduction. Cette typologie discerne aussi les formes ontologique et épistémologique et revient principalement à classer les caractérisations de la relation entre relata constituant la réduction. Voici succinctement en quoi elle consiste (voir Silberstein 2002 pour plus de détails).

Du côté ontologique, on retrouve d'abord l'*élimination* où X est remplacé par Y en sorte que X est éliminé de notre ontologie. Il y a ensuite l'*identité*, où X est identique à Y , non pas que X soit inexistant mais seulement que sa vraie nature serait révélée par Y . Puis viennent deux types de relations particulièrement importants pour le réductionnisme et conséquemment pour l'émergentisme. D'une part, la *survenance méréologique* stipule que les propriétés du tout sont déterminées par les propriétés intrinsèques des parties les plus fondamentales. D'autre part, la *survenance nomologique* stipule que les lois les plus fondamentales comme celles de la physique déterminent, par nécessité nomique, les lois des sciences dites spéciales, comme la biologie. Du côté épistémologique, on retrouve le *remplacement*, soit l'analogue de l'élimination pour les relata épistémologiques, comme les concepts et les théories. Ensuite, il y a la réduction *théorético-dérivationnelle* issue de l'approche orthodoxe du réductionnisme inter-théorique (présentée un peu plus loin). Une autre façon de caractériser une relation réductionniste de type épistémologique vient de l'approche *sémantique*, qui est somme toute révisionniste, puisqu'elle rejette plusieurs conceptions épistémologiques largement acceptées, comme celle voulant que les théories scientifiques soient des systèmes axiomatiques, ou encore que la logique de premier ordre soit adéquate à leur interprétation. Selon cette approche les théories scientifiques ne sont pas des entités linguistiques, des ensembles de phrases, mais des modèles mathématiques. Enfin, l'approche *pragmatique* soutient que si une théorie à un « niveau inférieur » offre une explication et une valeur prédictive jugées supérieures à ce que propose une théorie à un « niveau supérieur », alors il s'agit d'une réduction inter-théoriques réussie.

2.1.3.2 ORTHODOXIE ET CRITIQUES

Plusieurs explications qui paraissent tout à fait satisfaisantes en général, c'est-à-dire « dans la vie de tous les jours », endossées par l'intuition, sont de type réductionniste. Par exemple, l'explication visant à dire ce qu'est une horloge ou pourquoi une horloge est ce qu'elle est peut bien se résumer à une énumération de ces composantes, autrement dit en réduisant un système complexe en ses parties (plus) élémentaires. L'orthodoxie de l'explication scientifique a principalement, mais pas exclusivement, adopté une approche réductionniste. Les exemples abondent. La physique de Newton a su ramener la physique du monde terrestre et du monde céleste d'Aristote à quelques lois, Maxwell a montré que la lumière n'était « que » des ondes électromagnétiques, et maintenant on tente de réduire la mécanique

classique à la mécanique quantique, ou encore la génétique à la biochimie. Mais tous ces exemples, et il y en a bien d'autres, ne sont pas aussi simples qu'on voudrait bien le croire et cette entreprise de réduction d'une théorie à une autre constitue l'un des problèmes les plus « glissants » (« *slipperiest* », Sklar 1993 : 336) de la philosophie.

La conception orthodoxe d'une réduction épistémologique est due à Hempel & Oppenheim (1948) et Nagel (1961). L'idée de base est qu'une théorie, dite « réduite », T' , se réduit à une théorie, dite « réductrice », T , si et seulement si T' est dérivable ou déductible de T avec l'éventuel concours de « lois-ponts » (Nagel les appelle des « définitions de coordination »). Une telle réduction serait un certain type d'explication puisque la déductibilité d'une théorie par une autre exemplifie le modèle déductivo-nomologique. Mais le lien qui peut exister entre une théorie considérée comme fondamentale et une théorie à un niveau supérieur – l'exemple paradigmatique étant celui de la mécanique statistique et de la thermodynamique – est beaucoup plus complexe qu'on a pu le croire il y a quelques décennies (Sklar 1993 ; Primas 1998 ; van Riel & van Gulick 2014). Entre autres, le critère de déductibilité ne s'applique pas de manière aussi directe qu'il n'y paraît et cette conception d'inspiration positiviste n'est pas sans poser problème. Si l'interprétation des théories comme des systèmes hypothético-déductifs peut paraître acceptable, la caractérisation de la déductibilité est plus problématique.

Ce critère est souvent interprété comme l'exigence que la théorie réductrice puisse générer toutes les conséquences observables de la théorie réduite. Cette exigence est minimale et on s'attend à plus de la nouvelle théorie, qu'elle soit supérieure en un certain sens¹². Et elle peut être jugée supérieure de bien des façons. Elle peut l'être, entre autres, si elle rend compte de plus de phénomènes, autrement dit si elle *explique* davantage que la théorie réduite. En ce sens, elle serait plus générale. Toutefois, si la nouvelle théorie, réductrice, procède à une réinterprétation de phénomènes observables selon une unification conceptuelle, alors la généralité concerne non pas le type de phénomènes mais plutôt des phénomènes singuliers, spatio-temporellement localisés, bref des événements. Par exemple, la théorie électromagnétique ne fait pas de catégorie distincte pour la lumière et les ondes électromagnétiques, ces deux concepts se ramenant à un seul, en sorte qu'elle n'est pas plus générale en couvrant deux types de phénomènes, soit ceux de la lumière *et* ceux des ondes électromagnétiques mais bien, plutôt, de l'ensemble des phénomènes des ondes électromagnétiques *dont* ceux de la lumière. Dans ce cas, la

¹² On peut soutenir toutefois que cette exigence soit la seule qui compte vraiment. Elle serait donc non seulement nécessaire mais aussi suffisante. J'ai discuté plus tôt des exigences d'une explication, comme la cohérence interne et externe, qui s'applique ici aussi. Dans ce contexte, certains ont avancé que les « théories à variables cachées », dans le cas des phénomènes quantiques, n'apporteraient rien de plus aux théories sans variables cachées si elles génèrent les mêmes conséquences observables.

théorie réductive peut être jugée supérieure en étant plus simple, c'est-à-dire qu'elle rendrait compte des mêmes phénomènes avec moins d'appareils théoriques. Une théorie plus générale et plus simple aurait un plus grand « pouvoir systématique » (Sklar 1993 : 334). Mais ce qui se dégage de cette caractérisation sont les difficultés de vraiment déduire une théorie à partir de l'autre si ces deux théories n'ont pas les mêmes concepts, si les concepts des deux théories sont hétérogènes ou encore « autonomes » en ce qu'ils font partie d'un appareillage théorique satisfaisant certains critères épistémologiques comme expliquer certaines classes de phénomènes¹³. Dans la plupart des cas, ce qui est dérivé n'est pas exactement la théorie initiale, celle devant être réduite, mais plutôt une « image » de cette théorie au sein de la théorie réductive. Et cette image peut prendre plusieurs formes.

La traduction d'un certain concept d'une discipline scientifique vers une autre discipline, présentant des concepts différents, des « images », est susceptible de présenter des différences de signification, autrement dit de ne pas avoir la même intension, comme c'est le cas pour la traduction linguistique. La théorie réductive présente ainsi de nouveaux concepts par rapport à la théorie réduite – il y a là ce qu'on peut appeler une « discontinuité conceptuelle ». Tel que mentionné, ces concepts peuvent mener à l'élimination de certaines entités de notre catalogue ontologique ou encore au remplacement d'anciens concepts désormais jugés incorrects ou non pertinents. Ce fut le cas notamment de la possession démonique en psychologie et du phlogistique en chimie. La transformation peut aussi être moins drastique, c'est-à-dire que certains concepts de la théorie réduite peuvent demeurer pertinents d'un point de vue scientifique. La théorie réductive procède alors à une identification ontologique ou épistémologique où l'intension du concept est changée mais l'extension demeure la même. Par exemple, l'intension de « humain » n'est pas la même que celle de « bipède sans fourrure » mais présente néanmoins la même extension, ou encore la température d'un corps ne serait que l'énergie cinétique moyenne des molécules constituant ce corps.

Dans cette optique, Hempel (1965 : 103) affirme qu'une réduction de ce type consiste à « donner les conditions nécessaires et suffisantes dans la perspective de la théorie réductrice pour l'applicabilité des concepts de la théorie réduite ». La détermination de ces conditions est le fruit de recherches scientifiques bien souvent intenses et non d'une découverte logique ou philosophique, quoiqu'une restructuration taxonomique, comme dans le cas du « bipède sans fourrure » demeure possible. Par ailleurs, la pertinence de conserver certains concepts, sa justification, est généralement de deux ordres. D'une part, il peut s'agir de considérations cognitives ou représentationnelles, comme quoi les anciens

¹³ En ce sens, la biologie serait « autonome » de la physique puisqu'elle fournit des explications satisfaisantes. Dans certains cas, les sciences dites spéciales ou non-fondamentales pourraient fournir des explications plus « profondes », en ce sens où elles sont plus générales, et entendu que la généralité soit une valeur explicative, que les sciences comme la physique (Weslake 2010).

concepts permettent malgré tout une meilleure compréhension du phénomène, parce qu'ils sont plus simples ou plus intuitifs. Par exemple, le terme « respiration » n'entraîne pas la même compréhension sous des concepts intuitifs et n'a pas la même portée phénoménologique que l'expression « synthèse ATP », ou encore une pensée (un « qualia ») n'a pas la même signification ni la même représentation qu'une activité neuronale. D'autre part, et en lien avec la précédente justification, il se peut que les anciens concepts permettent une certaine appréhension empirique de ces concepts, qu'ils présentent un caractère observable ou mesurable. Par exemple, la température d'un corps macroscopique est mesurable empiriquement, contrairement à l'énergie cinétique des molécules.

L'image, discutée précédemment, censée représenter la théorie réduite, pose d'autres problèmes quant à la façon dont est obtenue cette « image », ce qui nous ramène à la caractérisation de la déductibilité du modèle théorético-dérivationnel. D'abord, et c'est pourquoi on parle d'« image », les déductions logiques d'énoncés exprimés exclusivement dans les termes de la théorie réductrice ne peuvent donner aucun énoncé caractéristique de la théorie réduite, car en toute rigueur logique ceux-ci devraient être présents dans les prémisses. Cependant, comme le souligne Hempel (1965 : 104), de manière triviale et arbitraire un nouvel énoncé peut être logiquement déduit d'un énoncé nomique sous forme conditionnelle ; il donne l'exemple de l'énoncé « lorsqu'un gaz est chauffé à pression constante, il prend de l'expansion », duquel on peut déduire « lorsqu'un gaz est chauffé à pression constante, il prend de l'expansion ou il se transforme en un essaim de moustiques ». En effet, puisque l'énoncé conditionnel, qui peut s'exprimer par une implication logique, est vrai, alors l'énoncé disjonctif, formé de ce conditionnel et de tout autre énoncé arbitraire (voire faux), est nécessairement et trivialement vrai. Bien que nécessaire, la déductibilité au sens stricte est donc insuffisante à la réduction épistémique¹⁴.

Ensuite, en raison de ces contraintes logiques, le passage théorético-dérivationnel d'une théorie à l'autre peut être rendu plausible par des « lois-ponts », mentionnées plus tôt. Celles-ci sont des propositions nomiques contenant les concepts des théories réductives et réduites, qui prennent la forme

¹⁴ Il est opportun de distinguer deux types de déductibilité (et de prédictibilité) dans le contexte d'une réduction épistémologique. Ce qui les distingue concerne la présence initiale ou non des propriétés théoriques dans les prémisses de l'inférence déductive censée opérer la réduction. Plus précisément, puisqu'un prédicat absent des prémisses ne peut en toute rigueur logique se trouver dans la conclusion, comme mentionné, il semble y avoir une différence notable entre une déductibilité où l'on cherche sciemment à obtenir une certaine propriété et une déductibilité ne reposant que sur certains principes fondamentaux. Ce semble être ce qu'entend van Cleve (1990, in McLaughlin 2008) par ce qui est « initialement prédictible » ou non. Par exemple, en mécanique statistique, l'irréversibilité définie initialement par l'augmentation d'entropie thermodynamique fait partie des desiderata de cette théorie réductive. Par contre, la proportionnalité inverse de la viscosité et de la température des gaz était totalement inconnue avant les travaux de Maxwell l'ayant mise en lumière. Ce n'est pas tant que la première déductibilité soit inopérante ou invalide, mais plutôt qu'elle semble, *prima facie*, ne pas entraîner le même degré de *confiance* en le résultat, à l'instar de la différence de confiance que l'on peut avoir entre un raisonnement visant à établir rétrospectivement, « en amont » si j'ose dire, les prémisses pour une conclusion particulière et un raisonnement plus conventionnel, « en aval ». Reste à voir si cette confiance peut se traduire en une certaine valeur épistémologique.

syntaxique de biconditionnels connectant les termes des deux théories. Mais la forme logique des lois-ponts n'est pas le souci premier. Plutôt, si la théorie réduite découle de la théorie réductrice, c'est-à-dire que la première est expliquée d'une certaine façon par la seconde par le truchement de ces lois-ponts, celles-ci ne devraient-elles pas être justifiées afin de jouer le rôle qu'on veut bien leur faire jouer ? Et si la théorie réduite est fausse, même partiellement, alors ces lois-ponts ne seraient que partiellement vraies, pour ne pas dire complètement fausses. La théorie faisant l'objet de la réduction risque alors d'être corrigée plutôt que récupérée par la théorie réductive, ou encore le gain explicatif pourrait être nul si de nouveaux éléments comme les lois-points demeurent injustifiés. De plus, dans le cadre d'analyses inter-théoriques, il devient difficile de bien cerner à quel cadre théorique se rattache ces lois-ponts, de sorte que les conséquences philosophiques qui les concernent demeurent floues : à quel niveau d'organisation appartiennent-elles, peut-on parler d'émergence de lois-ponts, etc. ? On peut se demander par exemple si l'entropie thermodynamique se réduit bel et bien à la mécanique si l'on doit faire appel aux probabilités qui n'en font pas partie.

Un autre problème de taille dans la conceptualisation d'une réduction inter-théorique concerne le holisme sémantique. Si la signification des termes et expressions d'une théorie est déterminée par le rôle qu'ils jouent au sein du réseau d'assertions de la théorie, alors il faut s'attendre à ce que tout changement, si minime soit-il, dans ce réseau puisse avoir des conséquences sur la signification de chacun des termes et expressions de cette théorie. Ainsi, les théories réductrice et réduite seraient totalement incompatibles. La thèse voulant que la théorie réduite soit une bonne « approximation » de la théorie réductrice serait alors compromise. Les exemples de limite asymptotique, comme la limite d'une célérité infinie dans la réduction de la mécanique newtonienne à la relativité restreinte ou la constante de Planck tendant vers zéro dans le cas de la réduction de la mécanique classique à la mécanique quantique, ou encore de la limite thermodynamique, où le nombre de molécules tend vers l'infini dans celui de la réduction de la thermodynamique à la mécanique statistique, sont beaucoup « trop subtils et trop complexes d'un point de vue formel » (Sklar 1993 : 336) pour recevoir un traitement unificateur. De plus, en opérant ce genre de procédure calculatoire, plus souvent qu'autrement on en vient à faire appel à des idéalizations dont le statut épistémologique est loin d'être certain. Par exemple, il n'est pas évident qu'une explication considérant des fictions comme des systèmes infinis puisse expliquer un phénomène observable ou encore une autre théorie (Jodoin 2011).

Enfin, un problème fréquent dans l'analyse conceptuelle de relations inter-théorétiques est celui de la réalisabilité multiple (section 2.2.5). Le problème est le suivant : la description d'un phénomène à un niveau « supérieur », découlant par exemple de la théorie réduite, est compatible avec plusieurs descriptions à un niveau « inférieur », autrement dit le phénomène « supérieur » est réalisable de différentes façons comme phénomène « inférieur ». Des exemples sont de mise : la douleur semble être

réalisable de plusieurs façons chez des animaux présentant des anatomies et des physiologies fort différentes ; deux systèmes peuvent être en équilibre l'un avec l'autre et avoir la même température mais être de nature très distincte quant à leur composition moléculaire ; le même programme informatique peut être réalisé ou opéré par des ordinateurs fort différents. Alors en quoi est-ce un problème ? En fait, plusieurs conceptions réductionnistes, qu'elles soient de type ontologique ou épistémologique, ont maille à partir avec la réalisabilité multiple car il devient difficile d'identifier une propriété particulière à un *ensemble disjonctif* de propriétés. Il faut aussi, dans ce cas, préciser en quoi les propriétés du tout sont déterminées par celles des parties, ou encore comment le concept de la théorie réduite peut être dérivé de la théorie réductrice si cette dernière offre un ensemble disjonctif formé des multiples possibilités compatibles avec ledit concept.

Un plus grand pouvoir systématique peut bien entendu être gagné au prix d'une restructuration conceptuelle, mais si l'exigence de rendre compte, et peut-être même d'expliquer, plus de phénomènes est respectée, nonobstant les difficultés susmentionnées, l'ajout conceptuel peut être jugé problématique sous d'autres motifs. Cet ajout peut, par exemple, concerner des entités inobservables (discuté plus loin), qui pour diverses raisons peuvent être jugées comme douteuses. C'est le cas notamment des atomes, à l'époque, dans la théorie cinétique ou encore, plus récemment, des « supercordes » ou de certaines particules dites exotiques dans les tentatives d'unification de la physique des particules, de la mécanique quantique et de la relativité générale. Ce genre de difficultés, allant au-delà d'un simple manque d'imagination, peut mener à la conclusion selon laquelle la discipline contenant l'appareil théorique ne pouvant être ainsi réduit est autonome, selon des critères généralement liés à l'explication.

Une réduction épistémologique est donc, minimalement, une reformulation conceptuelle. Mais il faut, bien sûr, davantage. La question est alors de spécifier correctement les exigences, souvent formelles, de la réduction, mais la dérivabilité pose des problèmes d'ordre logique que les lois-ponts ne peuvent complètement dissoudre. Même si ces problèmes pouvaient être résolus, des obstacles demeurerait, à tout le moins des perplexités, d'ordre sémantique. Car on peut mitiger ces exigences formelles en s'appuyant sur le critère d'explicabilité, qui demande tout de même une clarification de ce qui est entendu par cette notion polysémique. Autrement dit, si l'on peut expliquer sans le critère strict de la déductibilité, alors on peut offrir une réduction sans déductibilité. Ce faisant, par contre, on risque d'offrir deux ensembles de concepts plus ou moins hétérogènes et sans doute incommensurables, de sorte que les vertus cognitives ou représentationnelles de l'un et de l'autre ne soient pas coextensives. Il est alors tentant, comme dans le cas de la traduction linguistique, de conserver ces deux ensembles de concepts. Réside donc un doute, une perplexité, sur la réelle réussite de la réduction. En conséquence, les réductions totales, comme dans les cas d'une élimination ou une identité, sont plutôt rares et la plupart des explications font appel à des hypothèses à plusieurs niveaux d'organisation (section 2.2)

différents. Car si l'exposition des parties constituantes d'un tout par l'analyse (une forme de réduction donc) est un excellent moyen de le comprendre et sans doute de l'expliquer, les régularités (causales ou non) qu'il manifeste sont justement, bien souvent, à un niveau où la description de son comportement n'est pas rendu plus intelligible ou compréhensible par un appel à la description de ses parties constituantes.

2.1.4 Un modèle d'explication

On constate qu'il y a plusieurs modèles d'explication, plusieurs théorisations en philosophie, plusieurs stratégies d'explications en sciences. On le répète, ce pluralisme ne devrait toutefois pas cautionner un relativisme. Bien que plusieurs objectifs, apparemment antagonistes, puissent être poursuivis avec une explication, tout n'explique pas pour autant. Il convient donc d'identifier clairement les éléments constitutifs de ce qui peut être légitimement considéré comme une bonne explication. En général, si une prétendue explication fournit une « expectabilité » suffisante, c'est-à-dire si l'explanans fait en sorte que l'on soit en droit de s'attendre à l'occurrence du phénomène décrit par l'explanandum, alors on juge l'explication satisfaisante. Précisons.

Dans le modèle déductivo-nomologique, « la raison selon laquelle un argument déductif valide quant à l'occurrence d'un certain événement explique cet événement est précisément que cet argument justifie, dans la mesure où les prémisses sont justifiées, l'expectative que l'événement se produira »¹⁵ (Weslake 2010 : 274-5). Et bien évidemment, la causalité offre aussi cette relation de dépendance exprimée par un conditionnel contrefactuel. Ainsi, Mill (1882 : 402) soutient que si une personne venait à manger un repas et à en mourir, c'est-à-dire qu'elle ne serait pas morte si elle ne l'avait pas mangé, alors on serait en droit d'affirmer que le repas est la cause de la mort. Il en va de même d'un mécanisme comme ensemble de processus « produisant certains résultats de manière régulière » (Machamer, Darden & Craven 2000 : 7). L'orthodoxie de l'explication scientifique satisfait donc le critère d'expectabilité.

Mais ce critère, issu d'abord du modèle déductivo-nomologique, a été critiqué non pas sur sa pertinence ou son caractère nécessaire, mais par rapport à sa suffisance, comme quoi il manquerait quelque chose pour bien cerner la notion d'explication. L'une de ces critiques est celle des asymétries explicatives (discutées plus tôt) : dans le cas d'une loi exprimant des corrélations entre différentes

¹⁵ « [...] the reason that a sound deductive argument for the occurrence of some event explains that event is precisely that the argument justifies, to the extent that the premises are justified, the expectation that the event will (or did) occur ».

variables, il est possible de dériver la valeur d'une d'entre elles, et donc d'exemplifier l'expectabilité par la déduction, même si cette dérivation ne semble avoir de pouvoir explicatif. Par exemple, on peut dériver la hauteur d'un bâtiment à partir de l'ombre qu'il projette, mais on peut douter que l'ombre explique la hauteur. Une autre est celle des impertinences explicatives : il est possible de déduire l'explanandum, à partir d'une généralisation universelle supportant les contrefactuels, qui semble impertinente ou inexplicative de l'explanandum. Il a été mentionné qu'une stratégie notoire pour répondre à ces critiques est d'insister sur les relations causales (par ex., la hauteur cause l'ombre, mais pas l'inverse, et la prise de pilule contraceptive par un homme ne cause pas sa non-grossesse). Pour Woodward (2003a : 191), une explication exige la démonstration de « schèmes systématiques de dépendance contrefactuelle » (« *systematic patterns of counterfactual dependence* »)¹⁶. Il s'agit en substance de fournir des réponses à des questions contrefactuelles, de dire en quoi l'explanandum serait différent si les choses avaient été différentes. Cette approche est intéressante puisqu'elle permet de rendre compte de plusieurs des objectifs généralement visés avec une explication et donc du pluralisme (restreint) que rencontre la notion d'explication. Bien entendu, ces schèmes systématiques sont déterminés par des interventions selon la stratégie M (section 2.1.1). Ce critère supplémentaire permet de pallier, dans une certaine mesure, aux objections précédemment soulevées. Puisque l'insuffisance d'un concept peut être comblée par sa réinterprétation, tentons maintenant de cerner davantage ce que signifie l'expectabilité.

L'expectabilité peut être, en effet, caractérisée de plusieurs façons. En un sens, elle possède une connotation psychologique, presque subjective. Bien entendu, la notion d'explication est liée de près à celle de compréhension : si l'on m'explique pourquoi il y a des arcs-en-ciel, je serais sans doute en mesure de dire que je comprends pourquoi et comment il y a des arcs-en-ciel. Mais l'expectabilité dans son acception psychologique peut être considérée non pas comme une condition d'une bonne explication mais plutôt comme sa conséquence. Je propose ainsi deux façons de caractériser une explication, pouvant soutenir une expectabilité et ces schèmes systématiques, qui ne sont pas mutuellement exclusives mais complémentaires. C'est-à-dire que leur extensivité, leur « domaine de juridiction », ne sera pas clairement définie selon les différents types d'explication. Mais elles permettent tout de même de circonscrire davantage les exigences d'une explication satisfaisante. De plus, elles permettent de rendre compte du pluralisme explicatif rencontré dans les systèmes complexes où plusieurs stratégies explicatives, à plusieurs « niveaux » différents, dans plusieurs disciplines différentes, concourent à fournir une compréhension de leur comportement. Il suffit de penser, dans le cas des traits des organismes, aux différentes stratégies explicatives de l'écologie, de la génétique des populations et de la biologie développementale.

¹⁶ Thalos (2002) soutient plutôt qu'une explication est une « *illumination of a kind of dependence relation* ».

2.1.4.1 MODALITÉ

La première est la « modalité », selon laquelle une explication doit montrer que ce qui s'est passé, et qui doit être expliqué, « devait » se produire. Toute la difficulté de ce critère se situe dans la justification de la force modale de ce « devait » (était-ce nécessaire, etc. ?). Une force trop grande le commet au déterminisme absolu, qui peut être d'emblée écarté¹⁷, tandis qu'une force trop faible autorise une trop large gamme d'explanans et le confine à la vacuité. En se situant entre ces deux extrêmes, la modalité doit garantir un espace conceptuel suffisant pour les explications probabilistes, avec des lois statistiques, mais aussi pour l'explication d'occurrence particulière. L'exemple de l'explication de la couleur bleue des yeux d'un certain enfant, qui n'est en rien « nécessaire », semble montrer que le critère modal, au sens fort, est intenable. Ce qui est nécessaire ce sont les proportions statistiques de yeux bleus et bruns, pour une population suffisamment large, déterminées par les lois de Mendel. Mais, justement, la modalité n'a pas besoin de signifier une nécessité au sens fort, logique du terme (Grimm 2008). Elle peut, par exemple, signifier quelque chose comme « cela avait de fortes chances de se produire » plutôt que « cela devait se produire ». En interprétant la modalité de la sorte, c'est-à-dire en se référant à des concepts comme celui de chance ou de probabilité, c'est au concept de possible que l'on renvoie celui de modalité. Le concept de possible est déjà lié à celui d'explication car expliquer un phénomène c'est dire que ce phénomène est possible. Expliquer pourquoi un phénomène est impossible n'est ainsi que l'opération logique inverse. Donc expliquer un phénomène implique de le situer dans une gamme de possibles. Cela répond, en un sens large, à une « question-pourquoi » du genre « pourquoi tel phénomène et pas un autre ? ». Dans l'exemple des yeux bleus, une explication se référant aux lois de Mendel n'offre pas une modalité au sens fort, mais présente déjà une gamme de possibles (des yeux bruns ou bleus), en opposition contrastive à une gamme d'impossibles (des yeux noirs, verts ou rouges, etc.), définissant une modalité au sens faible.

Le concept de gamme de possibles peut se représenter aisément par celui d'espace des phases, employé en mécanique statistique et en mécanique quantique (Chapitre 5). L'identification de cette gamme de possibles, la définition d'un espace des phases pour un système donné, dépend des propriétés que l'on attribue au système en question. Si le système est un cheval, par exemple, la gamme des possibles (espace des phases), naturellement, ne contient pas un cheval volant ou se trouvant à mille mètres de profondeur sous la mer. Cela permet de définir deux classes d'énoncés contrastives définissant ce qui est possible en opposition à ce qui ne l'est pas (et d'offrir des explications élémentaires comme « le cheval ne vole pas parce que ce n'est pas un oiseau »). Mais une explication exige

¹⁷ Il n'est pas tant question d'éliminer la *possibilité métaphysique* du déterminisme absolu que d'écarter l'*exigence épistémique* qu'il pourrait placer sur l'explication.

davantage : il faut circonscrire l'espace des phases ou identifier l'une de ses régions. Dans un cas déterministe comme celui de la mécanique classique, cela revient à déterminer une trajectoire, donc une suite de points, dans l'espace des phases. Les causes ou les lois sont censées opérer de la sorte. Une loi stricte offre évidemment ce caractère modal en connectant logiquement l'explanans à l'explanandum. Une cause est aussi, souvent, définie comme une condition nécessaire et suffisante à l'effet. C'est pourquoi le niveau de précision de la trajectoire dans l'espace des phases, dans cette gamme de possibles, détermine le niveau de précision de l'explication (l'explication du cheval qui ne vole pas ne serait *pas* précise).

Mais ce n'est pas si simple. Car définir une trajectoire dans l'espace des phases revient non seulement à présenter les conditions nécessaires et suffisantes au phénomène, mais aussi tous les détails microscopiques selon une évolution déterministe. Or, tel que mentionné, toutes les lois scientifiques ne sont pas déterministes comme en mécanique classique et il n'est pas certain que ces détails microscopiques soient accessibles si l'on en croit la mécanique quantique, ni même nécessaires à l'explication si l'on en croit la mécanique statistique. Autrement dit, il y a lieu de douter que l'explication ne soit possible que dans les cas déterministes. En outre, définir une cause selon des conditions nécessaires et suffisantes est très problématique pour plusieurs raisons. D'abord, des conditions nécessaires et suffisantes semblent, comme un biconditionnel, éliminer l'asymétrie entre la cause et l'effet. Ensuite, selon les différentes définitions de la causalité, les arguments de préemption et des régularités fallacieuses posent de sérieuses difficultés à la nécessité d'une cause pour son effet, tandis que les arguments de l'inaboutissement et de la méconnexion sont problématiques pour la suffisance (Schaffer 2007 ; Jodoin 2010a, 2010b). Enfin, et surtout, la nécessité et la suffisance d'une cause pour son effet sont en conflit l'une avec l'autre : une cause nécessaire doit être suffisamment « large » pour englober toutes les possibilités (pour s'assurer que le flot inverse, de l'effet vers la cause, dans l'espace des phases passe effectivement par la cause), alors qu'une cause suffisante doit être assez « étroite » pour inclure tous les détails de la cause (Eckhardt 2006).

Une modalité interprétée selon la causalité ou la nomicité stricte impose donc des conditions formelles trop fortes à l'explication. L'inclusion de l'explanandum dans une gamme de possibles, contrastée donc avec un ensemble d'états impossibles, pouvant soutenir les contrefactuels semble alors une approche susceptible d'éviter les problèmes propres à l'orthodoxie de l'explication. Il peut y avoir d'autres moyens d'appuyer de manière modale une trajectoire ou un ensemble de trajectoires dans un espace des phases, même si d'ordinaire on y réfère comme étant une « cause ». Les interventions de la stratégie M permettent, en établissant un schème systématique de dépendance contrefactuelle entre deux variables, de connecter de manière modale deux points de l'espace des phases, l'explanans et l'explanandum.

2.1.4.2 NATURALITÉ

La seconde interprétation de l'expectabilité visant à caractériser une explication est la « naturalité ». Workman (1964) soutient en effet qu'une caractéristique nécessaire à l'explication est que celle-ci doit conférer une « naturalité » (« *naturalness* »), ce qui confère un aspect naturel, à l'explanandum. Toutefois, le terme « naturel » a cette particularité de recouvrir sous les apparences du consensus les oppositions les plus tranchées. En un sens assez intuitif, une chose paraissant normal ou naturel ne semble pas exiger (plus) d'explication. Par exemple, Healey (1994, n20) affirme que les « idéaux d'un ordre naturel » (« *ideals of natural order* ») sont des « explications fondamentales » (« *bedrock explanations* »). En outre, ce qui est naturel est en général compris. Mais au-delà de ces acceptations relativement élémentaires de ce qu'on peut entendre par naturalité, qui au demeurant peuvent apporter un éclairage supplémentaire à cette notion fondamentale et primitive qu'est l'explication, ce qui importe est plutôt la manière dont est conféré cette propriété à de prétendues explications.

Selon l'auteur, cette naturalité peut être conférée par la subsumption d'un concept sous une classe de référents similaires (Workman parle de « choses » (« *things* »)) ; par la familiarité, soit par association avec quelque chose de connu ; par l'autorité (ce qui s'apparente davantage à une justification) ; par association à un « idéal d'un ordre naturel ». Il donne à titre d'exemple pour le dernier cas le principe newtonien d'inertie où il est « naturel » pour un corps de se mouvoir selon une trajectoire uniforme et rectiligne. L'explication de la chute des corps dans la physique aristotélicienne fait référence à cet « état naturel » qu'est le centre de la terre, le centre de l'univers vers lequel tendent les corps. En ce sens, la naturalité répond à l'exigence de la dépendance contrefactuelle où des circonstances différentes, des états différents, mènent au même résultat. En montrant que plusieurs conditions initiales d'un système évoluent vers un certain état d'équilibre, celui que l'on souhaite expliquer, les explications d'équilibre exemplifient aussi la naturalité. Les cas où est exemplifiée la réalisabilité multiple, où plusieurs détails microscopiques sont impertinents ou ne font pas de différence au niveau macroscopique, ne sont pas différents à cet égard. Par exemple, des systèmes dynamiques similaires et pourtant différents possédant un même attracteur évoluent donc vers un même état final caractérisé par cet attracteur. De tels systèmes n'étant pas purement chaotiques sont d'ailleurs considérés « naturels » (Kaufman 1993).

Par ailleurs, la naturalité, par la subsumption d'un concept sous une classe de référents similaires ou familiers, permet d'éliminer, à l'instar d'une gamme de possibles, les explanans incompatibles avec un certain arrière-plan ontologique, faisant intervenir, par exemple, des divinités ou des êtres invisibles. Bien entendu, le domaine de validité d'une loi circonscrit aussi, partiellement du moins et sans nécessairement faire référence à un arrière-plan ontologique, les entités pertinentes à l'explication. Un autre exemple pouvant être interprété par le critère de naturalité est celui d'une explication où il est

montré que l'explanandum est un cas particulier, un exemple, d'une loi, sans pourtant le déduire (Forge 1980).

Selon Salmon (1989), la nomicité, la modalité et la dépendance contrefactuelle semblent posséder une extension commune, c'est-à-dire qu'elles semblent s'appliquer aux mêmes énoncés. C'est la raison pour laquelle la modalité et la naturalité ont été décrits comme étant complémentaires¹⁸. Je ne puis élaborer davantage sur ces concepts : maintes précisions supplémentaires devraient être apportées afin de clarifier des concepts aussi généraux et parfois vagues que ceux de modalité et de naturalité. Ils semblent néanmoins offrir un cadre suffisamment bien défini à la conceptualisation d'une explication satisfaisante et suffisamment large pour tenir compte de son pluralisme. Le modèle d'explication proposé ici suggère ainsi d'emprunter la notion d'espace des phases afin de circonscrire la gamme des possibles pour un explanandum donné. En identifiant d'abord une région de l'espace des phases, sont définies deux classes d'énoncés contrastives circonscrivant ce qui est possible et ce qui ne l'est pas. Ensuite il faut appuyer un lien modal entre l'explanans et l'explanandum. Les critères pertinents de l'orthodoxie d'une « bonne » explication, comme les lois, les causes et les régularités statistiques permettent, dans une certaine mesure, de soutenir une expectabilité, de répondre à ces schèmes systématiques de dépendance contrefactuelle, mais exemplifient aussi les concepts de modalité (au sens faible) et de naturalité. Une « bonne » explication doit donc montrer que ce qui s'est passé (explanandum) pouvait (gamme de possibles) se passer, plutôt que tel autre explanandum (impossible) ; ou que cela devait, selon une modalité faible, se passer ; ou que dans des circonstances différentes cela se serait aussi, naturellement, produit.

¹⁸ Par exemple, Wong (2006 : 349, n9) affirme de la survenance : « *supervenience is not a “natural” relation but merely captures patterns of modal dependent variation* ».

2.2 Niveaux d'organisation

Les explications réductives se présentent généralement d'après une relation entre un tout et ses parties, en termes méréologique donc, selon au moins deux « niveaux ». La conceptualisation d'une hiérarchie de niveaux pose immédiatement une série de questions à propos de la *réalité* de ces niveaux et de leur traitement épistémique : sont-ils réels ou non, et en quel sens, selon quel critère faut-il les définir, se ramènent-ils à un seul niveau fondamental et « plus réel », et justement, la réalité peut-elle venir en degrés, et enfin la différence entre niveaux est-elle autre chose qu'une différence d'accessibilité épistémique, surtout d'accessibilité empirique ? L'approche fondationnaliste dans son acception épistémologique stipule que toute connaissance repose ultimement sur certaines connaissances non-inférentielles, c'est-à-dire qu'elles ne seraient pas inférées de connaissances plus fondamentales ou primitives ou simplement postulées, lesquelles sont généralement caractérisées selon des critères d'évidence ou d'indubitabilité. Sa contrepartie ontologique stipule que les entités auxquelles réfèrent ces connaissances non-inférentielles sont réelles. Dans l'interprétation d'une théorie comme système hypothético-déductif, tout l'édifice théorique s'appuie en définitive sur ces bases prétendument inébranlables. En cela il s'agit d'un « fardeau épistémologique » imposant, car si les bases flanchent, le reste de la théorie aussi. Il peut alors être opportun de s'éloigner d'une approche fondationnaliste, sans toutefois la rejeter complètement, car une structure hypothético-déductive axiomatisée peut bien entendu présenter des vertus heuristiques et sémantiques.

2.2.1 Un cadre réaliste

Wimsatt (1981 ; 1994) offre un « critère de réalité » en termes de *robustesse* – parfois aussi appelé *détermination multiple* – qui se démarque des critères d'évidence ou d'indubitabilité, sans toutefois les rejeter complètement. Selon lui, une entité, qu'elle soit concrète ou abstraite, est robuste si elle est accessible – c'est-à-dire détectable, mesurable, dérivable, définissable, prédictible, etc. – d'une multitude de manières indépendantes, par différents « modes d'accès ». Et si elle est robuste, elle est prétendument réelle. Nous aurions confiance dans les objets, propriétés, relations, etc., qui sont détectables, mesurables, dérivables, etc., d'une multitude de manières indépendantes, car la « probabilité » d'être simultanément dans l'erreur pour chacune de ces manières diminue avec le nombre de vérifications ou de moyens de détections indépendants. Par exemple, les objets ordinaires, comme une table, ont des frontières assez bien définies et peuvent être détectés de maintes façons (visuellement, tactilement, etc.) qui, en gros, coïncident. C'est-à-dire que les résultats et les

interprétations des différents modes d'accès convergent vers une même description de l'objet. Un mode d'accès valide peut aussi être rationnel ou formel, comme une dérivation mathématique ou logique.

En ce sens, ce critère de réalité ne se limite pas aux entités observables et englobe celui d'Einstein, Podolsky et Rosen (1935) : si, sans perturber d'aucune façon un système, l'on peut prédire avec certitude (c'est-à-dire avec une probabilité égale à l'unité) la valeur d'une grandeur physique, alors il existe un élément de la réalité physique correspondant à cette grandeur physique. Il s'agit donc d'un critère qui s'inscrit dans la tradition réaliste – une tradition qu'Einstein a d'ailleurs largement contribué à bâtir. Il autorise donc, contrairement à une approche d'obédience strictement empiriste, les inférences à la réalité d'une entité à partir d'un mode d'accès plus indirect et discursif faisant appel à une certaine rationalité (comme une preuve mathématique), bref à partir de preuves qui ne sont pas obligatoirement observationnelles. Il faut toutefois noter deux points importants. D'abord, le critère de réalité d'Einstein, Podolsky et Rosen n'est pas une condition nécessaire à celui de Wimsatt, de sorte que si le premier devait échouer le second ne serait pas nécessairement invalide, puisque la prédictibilité est un mode d'accès parmi d'autres. Ensuite, le critère de Wimsatt n'implique pas la certitude – il n'y a pas de « formule magique » (Wimsatt 1994 : 10) en sciences comme ailleurs (discuté plus loin).

La thèse générique du réalisme stipule (i) que le monde existe indépendamment de l'esprit humain et (ii) qu'il est connaissable au moins en partie (Bunge 2006 ; Kuipers 2007). Si l'on se commit à cette thèse, alors il est possible, quoique discutable, d'inférer des conclusions ontologiques, sur les atomes ou l'émergence par exemple, à partir d'hypothèses épistémologiques, tirées de nos meilleures théories scientifiques par exemple. En effet, si nos connaissances scientifiques « lèvent le voile » sur le monde, qu'elles offrent une certaine correspondance avec celui-ci, alors il est justifié d'effectuer ce genre d'inférence. L'objection la plus courante contre le réalisme scientifique vise l'apparente incongruité entre la discontinuité historique du savoir scientifique et la prétendue uniformité diachronique de la nature, en sorte que des théories incompatibles ne pourraient être toutes deux « vraies », même à des moments différents. Puisque, toujours selon cette objection, les théories actuelles n'offrent pas plus de garantie de vérité que leurs contreparties historiques qui ont été falsifiées ou simplement remplacées, le lien entre théorie scientifique et vérité serait brisé, contredisant du même coup la thèse réaliste. Ainsi, une conception ontologique inspirée de la physique classique aura maille à partir avec les développements de la physique quantique. C'est pourquoi l'on a tenté d'éviter une correspondance terme à terme entre les entités théoriques et les constituants individuels de cette « réalité extérieure » en proposant, entre autres, un « réalisme structural », une forme atténuée du réalisme scientifique qui s'abstient de (vouloir) lier son sort à celui d'une affirmation d'existence des objets supposés par la théorie (voir Bitbol 1998 ; Chakravartty 2004 et 2006 ; Soler 2006).

Mais, encore une fois, le réalisme scientifique n'est pas tenu à l'infailibilité des théories scientifiques. Et ce n'est pas parce que nos connaissances sont faillibles que le monde est inconnaissable : le réalisme dit que la connaissance de la réalité est possible, pas que nous ayons effectivement cette connaissance en main. Plutôt, la *possibilité* que nos descriptions de la réalité soient vraies n'implique pas non plus qu'il soit possible de *reconnaître* cette vérité. Autrement dit, une *définition* de la vérité n'implique pas un critère d'*identification* de la vérité. Dire qu'une théorie est vraie et une autre fausse consiste essentiellement à comparer deux ensembles de justifications, dont l'un paraît meilleur que l'autre, pas à reconnaître directement la vérité de l'une de ces théories. Ainsi, même si une telle description (croyance) reçoit des justifications qui semblent valides, aucun ensemble de justifications n'est suffisant pour offrir une garantie reconnaissable de sa vérité. En conséquence, nos connaissances ne doivent pas, puisqu'elles ne le peuvent pas, être infailibles, immuables, pour soutenir la thèse réaliste. (Pour plus de détails, voir Jodoin 2014a.)

Nul besoin donc d'endosser la thèse à l'effet que le réalisme ontologique n'a pas le « droit à l'erreur » en décrivant « ce qu'est » la réalité. L'argument appuyant cette thèse soutient que si les connaissances servant à former une catégorie ontologique sont faillibles, alors la thèse voulant que la réalité soit connaissable serait infondée. Or, toujours selon cet argument, puisque nos connaissances scientifiques sont faillibles et qu'elles constituent nos meilleures connaissances, alors la réalité, au sens ontologique du terme, serait inconnaissable. Mais pourquoi lier ainsi infailibilité et réalisme ? Pourquoi la possibilité (épistémique) de l'erreur impliquerait l'impossibilité de fournir une description valable de la réalité (ontologique) ? Si la métaphysique est une discipline, pourquoi ce faillibilisme lui serait fatal, contrairement à d'autres disciplines, comme la physique par exemple ? Même si toutes nos connaissances sont faillibles, il semble *a priori* permis de systématiser nos connaissances au sein d'une théorie ainsi unifiée, générale, que l'on campe généralement dans le champ de la métaphysique. Cette approche tient donc du naturalisme, comme quoi les hypothèses métaphysiques devraient être « dérivées » des découvertes et arguments scientifiques (Ross & Spurrett 2007).

Mais affirmer que certains outils de la métaphysique peuvent être utiles et que certaines questions qu'elle aborde et réponses qu'elle propose soient légitimes ne signifie pas qu'elle soit nécessaire à la philosophie des sciences ni au présent débat. Est donc récusée ici la thèse de Carnap à l'effet que les énoncés métaphysiques ne peuvent être « reconstruits rationnellement », car, entre autres, la testabilité n'est pas une condition nécessaire à la signification, mais plutôt l'inverse (Manher 2007). Physique et métaphysique s'enrichissent mutuellement (Belot & Earman 1999).

2.2.2 Conceptualisation des niveaux

Le concept de mode d'accès, dans son acception empirique du moins, peut être interprété selon les conceptions *interventionniste* (ou *manipulationniste*) de la réalité et de la causalité (discuté précédemment, section 2.1.1). La distinction entre une manipulation et une intervention n'est pas toujours simple ou consensuelle, mais, de manière générale, une manipulation est considérée comme ce qui permet un changement de valeur d'une variable ciblée, tandis qu'une intervention serait une manipulation présentant la caractéristique de *modularité*, c'est-à-dire un découplage de la variable modifiée par rapport aux autres variables indépendantes de cette même variable ; autrement dit, un changement de valeur de la variable n'affecterait pas celles des autres. Ce qui est recherché en définitive ce sont des éléments pouvant être articulés au sein d'une explication, où ce qui est considéré comme « réel » ou « causal », bref ce qui est épistémologiquement fondé, aurait un pouvoir explicatif. L'idée générale derrière ces conceptions remonte ainsi à la fonction même d'une explication entendue comme moyen de contrôle, contrairement à ceux permettant une compréhension ou une prédiction ou tout autre élément possédant une quelconque valeur épistémique. Ainsi, le contrôle, même temporaire, d'une entité par une certaine intervention augmenterait la croyance en l'existence de cette entité, ou des phénomènes où elle se manifeste au sein d'une relation causale. En établissant une corrélation entre variables selon la stratégie M, les interventions constituent aussi une heuristique importante des théories scientifiques et un moyen de les tester.

Or, un mode d'accès empirique, entendu comme une intervention, est relatif et spécifique à un certain niveau. Par exemple, un microscope électronique et une lampe de poche ne donnent pas accès à des entités au même niveau, et ces entités sont généralement décrits et leur comportement expliqué par des théories qui leur sont spécifiques. On dit alors qu'elles sont à des *niveaux d'organisation* différents. La conceptualisation de niveau d'organisation permet de définir une hiérarchie de niveau, qui peut être interprétée ontologiquement ou non, mais qui peut aussi servir d'éléments de base à une conception particulière de l'émergence (section 2.3). Une discussion des thèses de Wimsatt (1994 et 2012) et de Bunge (2003) permettra de circonscrire les éléments d'une définition de ce qu'est un niveau d'organisation. D'abord, chez ces auteurs, il est question avec une telle définition de caractériser l'organisation de la *matière* (dont la différence spécifique, pour Bunge, est la possession d'énergie). Ensuite, il importe de distinguer, dans un système en général, sa *composition* (collection des parties du système) de sa *structure* (collection de liens entre les parties). On peut au surplus définir l'*environnement* (collection d'items, autres que celles du système, interagissant avec certaines parties du système) et le *mécanisme* (collection des processus faisant que le système se comporte de telle façon).

Dans un premier temps, selon Wimsatt (1994 et 2012), un niveau d'organisation est un niveau compositionnel où les « choses » (« *stuff* ») sont organisées en relations tout-parties de manière relative, c'est-à-dire qu'un tout peut être la partie d'un autre tout à un autre niveau. Les niveaux d'organisation représenteraient, selon lui, une « caractéristique fondamentale, non arbitraire et extrêmement importante de l'architecture ontologique de notre monde naturel » (1994 : 225). Ils sont constitués des classes d'entités présentant habituellement des tailles et des propriétés dynamiques comparables, interagissant typiquement et principalement entre entités de même classe. Ils présentent globalement un ensemble de phénomènes suffisamment réguliers pour former une « extensivité relativement close » (« *apparent rough closure* »), et non pas un amalgame chaotique de phénomènes inintelligibles. Par exemple, des disciplines scientifiques comme la thermodynamique et la sociologie peuvent prétendre décrire et expliquer certains phénomènes sans avoir recours à la physique quantique. C'est d'ailleurs pourquoi, selon Wimsatt, les théories scientifiques seraient exprimées par niveaux parce que les entités auxquelles elles réfèrent se présentent justement par niveaux. En effet, les interventions constituent une heuristique importante de l'élaboration des théories, mais il y en d'autres (voir Franklin 2007). Il s'agit donc d'une thèse forte se plaçant en amont des théories scientifiques, même si elle est motivée par une étude des théories scientifiques. Il n'y a donc pas de point de vue neutre, de « nulle part », offrant une assise fondationnaliste.

Dans un second temps, pour Bunge (2003), la notion de niveau d'organisation n'a pas de portée ontologique, car elle ne réfère pas à une chose ou une entité mais à une *collection* de choses, et par conséquent renvoie à des concepts, non à des entités concrètes. Cependant, Bunge ne prétend pas que ce qui est observé à des niveaux supérieurs, comme les phénomènes biologiques ou sociologiques, n'est que pure illusion, contrairement, par exemple, aux phénomènes physiques. Ce serait plutôt certaines *propriétés* d'entités concrètes qui seraient observées à des niveaux supérieurs, des propriétés émergentes. Pour autant, les conceptions de Wimsatt et Bunge ne sont pas incompatibles. Car, malgré la portée ontologique attribuée aux niveaux d'organisation par Wimsatt, il n'est pas nécessairement question de réifier ces niveaux et d'adopter une espèce de pluralisme des entités concrètes ou des substances. Parler d'interactions entre niveaux serait plutôt une « ellipse » (Bunge 2003 : 134). En effet, dans ces deux conceptions, il est question des propriétés d'entités se présentant à un certain niveau et non des propriétés du niveau lui-même. Ils peuvent ainsi être interprétés comme des « maxima locaux de régularité et de prédictibilité [...] de l'organisation de la matière » (Wimsatt 1994 : 238). C'est-à-dire comme une régularité et une prédictibilité des propriétés des entités concrètes révélant les propriétés de la composition et de la structure de certains niveaux. En ce sens, le critère voulant que les entités soient organisées en relations méréologique (tout-parties) ne paraît pas nécessaire, bien qu'il soit d'une aide à la représentation. Généralement, sauf exception, le contenu d'une théorie à un niveau supérieur est peu

influencé par la composition des niveaux très inférieurs (par ex. on peut imaginer que la thermodynamique soit la même que la matière soit continue ou non)¹⁹.

L'élément important ici est donc que les niveaux d'organisation sont définis par des théories scientifiques systématisant nos observations et nos hypothèses à propos de ce qui est considéré « robuste », et inféré comme « réel ». Or, corrélativement et d'un point de vue conceptuel, les théories se présentent aussi par niveaux car les phénomènes présentent des régularités qui se laissent plus facilement appréhendées et conceptualisées par niveaux. Les niveaux d'organisation sont donc caractérisés par des entités présentant des structures, des fonctions et des propriétés spécifiques qui, dans une mesure toute relative, en garantissent l'autonomie explicative. Cette autonomie est relative car, d'une part, des explications réductives demeurent possibles, en principe du moins, et d'autre part, des éléments externes sont souvent nécessaires à leur mise en œuvre au sein de prédiction et d'explication, comme des hypothèses et des concepts provenant d'autres théories, entre autres par rapport à l'environnement du système étudié, ou à des abstractions logico-mathématiques. En effet, les hiérarchies de niveaux d'organisation sont rarement figées et l'influence peut être autant descendante (par ex. une structure cristalline contraint le comportement de ses constituants atomiques) qu'ascendante (par ex. la dynamique orbitale des électrons contraignent les angles de la structure cristalline). Franchir un pas supplémentaire en attribuant à ces niveaux une portée ontologique demande bien sûr une justification. La métaphysique peut ainsi attendre – à moins qu'elle soit un outil méthodologique visant aussi à systématiser nos connaissances et à forger des modèles et des représentations, pouvant offrir des moyens heuristiques pertinents. Plus précisément, il peut être légitime de discuter rationnellement de ce que Popper (1935 : 252) appelle une cas « d'hypostase métaphysique d'une règle méthodologique bien justifiée » (comme la recherche de lois et de relations causales).

Ainsi, un niveau d'organisation est une description d'un ensemble de propriétés et de régularités suffisamment stable pour être déterminées de manière robuste selon les termes de Wimsatt. Cette conception des niveaux d'organisation est compatible, sans l'impliquer, avec une approche causaliste des propriétés et n'adopte pas nécessairement un fondationnalisme ontologique où des entités élémentaires à un niveau fondamental posséderaient des propriétés intrinsèques, tandis que des entités de niveau supérieur présenteraient plutôt des propriétés structurales. Conséquemment, un niveau d'organisation peut être considéré comme un « niveau de réalité », entendu qu'il s'agit des propriétés des entités que le constituant et qui se révèlent par des modes d'accès généralement empiriques, à partir

¹⁹ Il peut être utile alors de parler d'*échelle* (« *scale* ») plutôt que de niveau, comme le suggère Batterman (2002). Je vais toutefois conserver le terme « niveau », puisqu'il est d'usage dans la littérature, et il ne semble pas lié au critère méréologique.

desquels est décrit l'ensemble de ces propriétés interprétées par nos théories scientifiques, alors que ces théories sont aussi, en partie, constituées par le biais de ces modes d'accès.

2.2.3 Inobservables

Il a été vu qu'une définition d'une hiérarchie de niveaux pouvait recevoir différentes justifications. Le statut et la valeur de ces justifications varient selon la position épistémologique adoptée, généralement l'empirisme ou le réalisme. Le premier tente de « sauver les phénomènes », de décrire adéquatement ce qui est observable, tandis que le second s'autorise à inférer l'existence d'entités inobservables. La question de l'observabilité et de l'inobservabilité est donc celle de l'accessibilité empirique et de la légitimité de certaines justifications²⁰. Elle est particulièrement importante dans la définition des niveaux d'organisation qui, bien entendu, présentent divers degrés d'accessibilité empirique. Au final, c'est la légitimité de certaines explications scientifiques qui est en cause. À cet égard, l'exemple de la thermodynamique et de la mécanique statistique est probant puisque l'on tente, entre autres, d'expliquer des phénomènes observables, comme l'équilibre thermique, par des phénomènes inobservables, le comportement statistique de grands ensembles de molécules.

En corollaire vient la question de ce qui est observable *en pratique* ou *en principe*. D'abord, il est évident que des développements technologiques, qui sont contingents, peuvent étendre, en pratique, les possibilités de l'observation et donc de l'accessibilité empirique. Il n'y a qu'à penser au télescope ou au microscope, où les développements technologiques mais aussi théoriques ont permis de voir des entités de plus en plus éloignées et de plus en plus petites. Une inobservabilité en pratique est donc historiquement située. Elle n'apporte donc pas d'explication ou à tout le moins d'éclairage significatif sur le comportement de ces entités, sur les phénomènes, mais seulement, quoique ce soit déjà beaucoup mais impertinent ici, sur l'histoire de ces développements technologiques et théoriques. L'inobservabilité en principe est en revanche plus intéressante à la présente discussion. Elle peut découler d'un principe théorique stipulant que certaines entités, ou caractéristiques de ces entités, sont inobservables en raison de limites physiques particulières définissant les possibilités technologiques ou encore parce qu'il est jugé inutile ou *en pratique* impossible d'observer certaines caractéristiques physiques. En mécanique statistique, par exemple, l'inobservabilité des atomes est particulière. Elle est

²⁰ La question de la réalité des entités, propriétés, états, etc., inobservables est très ancienne et très générale. Car à bien des égards la réponse qu'on y apporte caractérise ce qui peut être appelé une « époque scientifique ». Par exemple, dans la tradition médiévale, les qualités « occultes » s'opposaient aux qualités « sensibles » et en somme référaient aux entités, propriétés, états, etc., inobservables (Hutchison 1982) ; et la révolution scientifique du XVII^e siècle semble offrir une approche où il serait permis d'obtenir une connaissance « scientifique » de ces entités, propriétés, états, etc., inobservables (Grant 1962).

d'abord historiquement située puisque les atomes, à l'époque, n'étaient pas observables en raison des moyens technologiques disponibles et il n'était pas possible de calculer la trajectoire de leur nombre gigantesque impliqué dans un système observable. Une impossibilité pratique donc. Cependant, malgré des avancées majeures en puissance d'observation et de calcul, cette inobservabilité est et va demeurer *postulée* par cette théorie. C'est-à-dire qu'il semble inutile ou non nécessaire de connaître dans le détail toutes les caractéristiques de la myriade d'atomes des systèmes macroscopiques, précisément parce que ce serait la mobilisation de caractéristiques statistiques sur de grands ensembles, et non pas de caractéristiques individuelles, qui serviraient à l'explication du comportement des systèmes macroscopiques. En cela il s'agit d'une inobservabilité de principe, que l'on pourrait qualifier de méthodologique. En mécanique quantique, l'inobservabilité semble plus forte, dans le cas des variables conjuguées, puisqu'elle découle d'une impossibilité théorique impliquant des limitations au processus d'observation vu comme un processus d'interaction « causale » entre l'observateur (ou l'appareil de mesure) et le système observé, modifiant significativement ce système.

Ensuite, il est clair que certains aspects ou caractéristiques d'une entité particulière peuvent être observables alors que d'autres ne le sont pas ; par exemple, la face cachée de la Lune peut être visible aux occupants d'une navette spatiale, ou encore la vitesse des molécules d'un gaz n'est pas observable bien qu'il soit possible de les détecter. Dans ce dernier cas, les modes d'accès sont plutôt indirects, à l'instar des ombres chinoises ; par exemple, une chambre à bulles ne permet pas de *voir* les particules, mais plutôt la trace qu'elles laissent, un peu à la manière d'un avion, mais dans ce cas on ne peut voir l'équivalent de l'avion au bout de la trace, comme le fait remarquer van Fraassen (1980 : 17). En outre, une théorie peut postuler, en principe, le caractère inobservable de certaines entités, tandis que pour une autre théorie ces entités seraient observables ; par exemple, la question de la structure de la matière est impertinente, en principe, dans le cadre de la thermodynamique, mais elle est tout à fait pertinente en biochimie. Enfin, il est aussi évident que la simple possibilité d'observer quelque chose n'implique pas que les circonstances soient toujours adéquates à son observation ; par exemple, les satellites de Jupiter sont observables au moyen d'un télescope, mais aussi à l'œil nu si l'on se trouve suffisamment proche.

Ainsi, pour qu'une chose soit observable, elle doit pouvoir être observée *lorsque les circonstances sont adéquates*. Ces circonstances sont doubles : ce sont celles, d'une part, des moyens empiriques d'observation et, d'autre part, de la conception théorique, et même ontologique, affectant l'interprétation des résultats d'observation. En effet, la thèse de la neutralité des énoncés scientifiques, à l'effet que les énoncés scientifiques ne dépendraient que de l'observation sans être contaminés par quelques présupposés que ce soit, est difficilement défendable. Ce n'est pas parce que deux personnes observent le même objet qu'elles en donneront nécessairement la même description. Ainsi,

l'inobservabilité que stipule une théorie peut découler d'une inaccessibilité empirique, soit comme une impossibilité technologique, soit comme une impossibilité théorique indépendante des moyens technologiques, ou encore d'une nécessité épistémologique, par exemple parce que la stratégie explicative prescrit une abstraction des détails que pourrait idéalement révéler une détermination complète du système. On peut en effet juger que l'inobservabilité a une certaine valeur explicative ou heuristique par rapport aux phénomènes observables ; par exemple, la dynamique moléculaire (inobservable) expliquerait les phénomènes thermodynamiques (observables), ou encore des caractéristiques génétiques (inobservables) expliqueraient certains phénomènes évolutionnaires (observables). Par conséquent, l'inobservabilité implique souvent une sous-détermination des propriétés du système qu'une théorie peut jusqu'à un certain point compenser. Évidemment, cette compensation dépend du degré d'empirisme adopté car une déduction n'équivaut pas nécessairement à une observation (la signification de cette équivalence pouvant être subtile toutefois), après tout on tente bien de corroborer empiriquement les théories. Une position se coupant des inobservables (entités, états, propriétés, etc.) adopte un certain agnosticisme, qui peut être avisé ou non, mais ce faisant elle se coupe aussi de la possibilité de certaines explications. Comme le souligne Bunge (1997), on peut difficilement imaginer ce qu'aurait été la science physique du XVIII^e siècle si Newton s'était abstenu de traiter d'inobservables comme la gravitation.

En revanche, un discours prétendant à une certaine valeur épistémologique, tant pour le réaliste que pour l'empiriste, devrait fournir les conditions de sa testabilité ou falsifiabilité, lesquelles reposent en fin de compte, même indirectement, sur des observables. Les raisons pour qualifier une chose d'inobservable sont donc nombreuses et relatives : elles dépendent d'un contexte expérimental, théorique et épistémologique qu'il faut préciser. Une position réaliste se garde la possibilité d'articuler un discours portant sur des inobservables au sein d'une stratégie explicative des phénomènes observables, ce que refuse un certain type d'empirisme. La question est ainsi de savoir si la prudence est de mise et si le prix à payer est acceptable selon les circonstances et les présupposés épistémologiques, et même ontologiques, qu'on est prêt à endosser.

2.2.4 Survenance

Une caractérisation de la relation entre niveaux d'organisation, célèbre mais problématique, est celle s'appuyant sur le concept de *survenance* (« *supervenience* ») : un ensemble de propriétés *A* « surviennent » d'un ensemble de propriétés *B* si et seulement s'il ne peut y avoir de différence en *A* sans qu'il y ait de différence en *B* (Teller 2009 ; McLaughlin & Bennett 2011). Une autre définition est celle de Butterfield

(2011) : soit un ensemble d'objets, O , duquel sont définis deux ensembles de propriétés, A et B ; alors A « survient » de B , si et seulement si deux objets dans O dont toutes les propriétés correspondent dans A , correspondent aussi dans B . Bien que ces définitions soient en termes de propriétés, il n'y a pas d'indication claire que d'autres relata puissent exemplifier ce concept. La survenance est ainsi un concept formel, modal : si l'on sait que les A surviennent sur les B , alors on sait que les B nécessite d'une certaine façon les A . Par exemple, les propriétés mentales « surviendraient » sur les propriétés physiques. La difficulté réside donc principalement dans la spécification et la justification de cette modalité, de cette nécessité. Cette modalité peut se présenter selon divers degrés en ce que l'impossibilité exprimée – « il ne peut y avoir » – peut signifier une impossibilité logique ou encore une impossibilité nomologique, par rapport aux lois de la nature ou par rapport aux lois scientifiques.

Bien que ce concept se soit avéré moins éclairant qu'on aurait pu le croire, sa présence dans littérature indique qu'il n'est sans doute pas totalement inutile. Pour Teller (2009), le concept de survenance offre un cadre général, en évitant les détails, à la compréhension de corrélation entre propriétés et fournit des informations sur la « structure » de cette corrélation. Par exemple, la deuxième loi de Newton, « $F = ma$ », indiquerait que la force survient, avec une nécessité nomologique, sur les valeurs de masse et d'accélération. En ce sens, la survenance permet d'affirmer une relation de détermination (aussi appelée « définissabilité implicite »), de composition, de réalisation ou de dépendance (principalement ontologique), entre certaines propriétés (importante pour le concept d'émergence, section 2.3), sans toutefois en préciser davantage la nature, sans aller au-delà de ce cadre général, sans préciser la modalité au sein de ce concept. En somme, dire que certaines propriétés macroscopiques surviennent sur des propriétés microscopiques revient en quelque sorte à dire que les secondes sont plus fondamentales que les premières. En conséquence, les propriétés macroscopiques « survenantes » pourraient être expliquées d'une certaine façon par les propriétés microscopiques sur lesquelles elles surviennent. Mais le terme de « structure » évoqué par Teller semble trop fort puisque la relation de détermination que pose le concept de survenance reste largement imprécisée (Witmer 2006).

En outre, la survenance est souvent considérée comme une condition nécessaire à la réduction. Cela est trivial dans le cas d'une réduction *qua* identité, en sorte que si A se réduit à B , il ne peut y avoir de différence en A sans qu'il y ait de différence en B . La relation du concept de survenance avec celui d'émergence, en revanche, n'est pas toujours claire, bien qu'il soit considéré comme moins engageant que ce dernier du point de vue métaphysique. Certains voient plutôt la survenance comme une condition nécessaire mais non suffisante à l'émergence, en sorte que l'émergence impliquerait la survenance (Kim 1999). Ainsi, McLaughlin (2008), modifiant une conception de l'émergence proposée par van Cleve, propose une définition de l'émergence exigeant la survenance, définie par une nécessité

nomologique mais excluant la nécessité logique. On constate donc que, paradoxalement, le concept de survenance se présente autant comme une condition nécessaire à la réduction qu'à l'émergence, souvent considérée comme l'échec d'une réduction. Mais il n'y a rien de vraiment mystérieux ou paradoxal car la survenance n'est qu'une façon d'insister sur la relation de dépendance, l'« enracinement » (« *goundedness* ») entre niveaux d'organisation, évidemment très présente au sein du concept de réduction, mais aussi au sein de celui d'émergence qui renferme toutefois une certaine notion d'autonomie entre ces niveaux (section 2.3).

2.2.5 Réalisabilité multiple

Jusqu'ici, divers éléments utiles à la conceptualisation de systèmes complexes entendus comme un ensemble de sous-systèmes à différents niveaux d'organisation ont été présentés. On a vu que la relation entre ces niveaux peut exemplifier le concept de survenance où une différence à l'un de ces niveaux entraîne, selon une modalité qu'il faut préciser, une différence à un autre niveau. Bien entendu, ce n'est pas l'unique façon de caractériser une relation méréologique. La relation de détermination peut ainsi être incomplète, en un sens, et il n'y a donc pas de correspondance terme à terme entre les ensembles de propriétés à différents niveaux, ou encore ces ensembles peuvent être dits hétérogènes. Cette caractérisation fait appel au concept de *réalisabilité multiple*, qui définit des entités ou des propriétés de niveau supérieur (n) étant réalisées de plusieurs façons différentes par des entités ou des propriétés de niveau inférieur ($n - 1$). En philosophie de l'esprit par exemple, elle exprime l'idée qu'un type mental (propriété, événement, etc.) peut être instancié ou réalisé par plusieurs types physiques (connexions neuronales). Il est aussi exemplifié en physique, où plusieurs états dynamiques des molécules peuvent être compatibles avec une propriété thermodynamique donnée, en biologie, où plusieurs combinaisons génotypiques peuvent être associées à un trait phénotypique (ces deux derniers exemples seront discutés plus loin), ou encore en sociologie, où une certaine propriété comme « être une démocratie » peut être réalisée de multiples façons (Sawyer 2004). Il semble ainsi que plusieurs systèmes complexes puissent présenter des régularités pouvant être décrites avec peu de variables malgré la multitude d'états caractéristiques de leurs composants et la grande variabilité de ces mêmes états. Peut-on alors en conclure que la réalisabilité multiple soit un « fait général de la nature » (Wimsatt 1994) ?

Wimsatt (1994, 2006 et 2012) soutient que la réalisabilité multiple est une *conséquence* de l'existence des niveaux d'organisation. Puisque les descriptions des entités et des propriétés aux niveaux supérieurs le sont de manière plus « grossière » (« *coarse-grained* »), il y a moins de descriptions de macro-états distinguables que de descriptions de micro-états distinguables. Les descriptions de niveaux inférieurs

(ou microscopiques) offriraient plus de détails mais seraient moins générales que celles de niveaux supérieurs (ou macroscopiques). La réalisabilité multiple serait donc définie (i) pour un nombre beaucoup plus grand d'états possibles et distinguables à un certain niveau (micro-états) que d'états à un niveau supérieur associé (macro-états), principalement en raison de la plus grande précision du vocabulaire employé pour décrire les micro-états, et (ii) de l'identité numérique du système de niveau supérieur avec le système de niveau inférieur ainsi décrit. De sorte qu'un état macroscopique (ou macro-état) du système à un niveau supérieur est associé à plusieurs états microscopiques (ou micro-états) possibles d'un système à un niveau inférieur. Un tout formé d'au moins deux parties à des niveaux différents est un cas de réalisabilité multiple si l'ensemble des états à un niveau, d'une part, et l'ensemble des états à l'autre niveau, d'autre part, sont en *relation un-plusieurs* (« *many-one* ») : une propriété P à un certain niveau n serait réalisée ou décrite par plusieurs propriétés à un niveau inférieur $n - 1$, tel que $P = \{A_1 \text{ ou } A_2 \text{ ou } A_3 \text{ ou } \dots A_i\}$. Cette relation est donc *antisymétrique*, puisqu'il y a toujours plus de propriétés au niveau inférieur qu'au niveau supérieur (ou inversement), et *disjonctive*, en raison du fait que les propriétés A_i jugées équivalentes ne peuvent être conjointement ou simultanément réalisées. (Voir aussi Kim 1992 ; Heil 1999 ; Gillett 2003.)

Il y aurait prétendument plusieurs conséquences et corollaires à la réalisabilité multiple (voir Sober 1999 ; Battermann 2000) : (i) une non-déductibilité, à partir de P , car on ne peut pas savoir quelle propriété A_i est instanciée en raison de la disjonction ; (ii) une acausalité, car une relation un-plusieurs ne peut être causale étant donné que les propriétés P et $\{A_1 \text{ ou } A_2 \text{ ou } A_3 \text{ ou } \dots A_i\}$ sont simultanées, synchroniques ; (iii) une non-explicabilité, puisque les propriétés de niveaux inférieurs ne peuvent pas expliquer les propriétés de niveaux supérieurs (P), parce qu'elles ne présentent pas une généralité suffisante, ou parce qu'elles ne peuvent satisfaire les critères d'une explication causale, ou parce qu'elles sont sans importance en regard du comportement global de P .

Si la non-déductibilité est bien une conséquence de la réalisabilité multiple, alors, selon la conception orthodoxe d'une réduction épistémologique, il ne faut s'étonner qu'on ait soutenu (par ex. Fodor 1968 et 1974) qu'elle implique l'irréductibilité. Cette dernière découlerait non pas du caractère disjonctif des propriétés au niveau inférieur (microscopique) réalisant une certaine propriété au niveau supérieur (macroscopique), mais plutôt de leur hétérogénéité. C'est-à-dire qu'il est à première vue impossible qu'il puisse y avoir des régularités macroscopiques « multiplement réalisées » par des propriétés microscopiques hétérogènes. Et une disjonction de types (« *kinds* ») hétérogènes ne serait pas elle-même un « vrai » type (Kim 1992). L'argument normatif veut aussi que ce qui ne constitue pas un vrai type ne puisse faire l'objet d'une véritable loi au sens de généralisation universelle supportant les contrefactuels. Afin de contrer cette objection à la réduction, d'autres (par ex. Walter 2006) ont soutenu que la disjonction des propriétés au niveau inférieur ne rendait pas la réduction impossible car elle

pouvait être reliée par des « lois-ponts », telles que : pour tout x , $Px \equiv (A_1x \text{ ou } A_2x \text{ ou } A_3x \text{ ou } \dots A_ix)$. Un contre-argument à cette stratégie fait aussi un appel normatif à la constitution des lois, et stipule que la disjonction ne peut représenter un « type scientifique », en sorte que les « lois-ponts » ne peuvent être des lois. À tout le moins, tel que discuté, les lois-ponts demande aussi une justification et celle-ci est d'autant plus nécessaire si elles doivent jouer un rôle explicatif.

Par conséquent, on peut défendre l'idée que la réalisabilité multiple implique l'émergence, ou qu'elle en est une condition nécessaire mais non suffisante, minimalement dans son acception épistémologique (section 2.3.1). De sorte que certaines disciplines scientifiques, comme la psychologie ou la biologie, ne pourraient être réduites à d'autres disciplines de niveau inférieur comme la physique. Sober (1999) soutient plutôt que la réalisabilité multiple ne mène pas nécessairement à l'échec de la réduction, sous une certaine forme du moins. Il avance que s'il y a bien asymétrie entre P et A_i , il n'y a pas nécessairement asymétrie entre P et l'ensemble disjonctif $\{A_1 \text{ ou } A_2 \text{ ou } A_3 \text{ ou } \dots A_i\}$. De sorte qu'une justification menant à l'affirmation que c'est le cas que P est équivalente à dire que c'est le cas que $\{A_1 \text{ ou } A_2 \text{ ou } A_3 \text{ ou } \dots A_i\}$, autrement dit que ces deux hypothèses sont confirmées de la même façon. En conséquence de quoi l'évolution dans le temps de l'ensemble disjonctif $\{A_1 \text{ ou } A_2 \text{ ou } A_3 \text{ ou } \dots A_i\}$ peut bien expliquer l'évolution dans le temps de P . Ce type d'explication est une microexplication diachronique (discuté précédemment). La non-explicabilité ne concernerait ainsi que les microexplications synchroniques mais pas les microexplications diachroniques.

À propos toujours de l'explication, selon Wimsatt (1994), une conséquence intéressante de la réalisabilité multiple est l'autonomie dynamique et (donc) l'autonomie explicative des variables et relations causales de niveaux supérieurs, car il y aurait stabilité relative des caractéristiques de niveaux supérieurs malgré les changements incessants aux niveaux inférieurs. En effet, tel que discuté, ceux-ci sont vus comme des maxima locaux de régularité et de prédictibilité alors que des changements au niveau inférieur n'entraînent pas nécessairement de changement au niveau supérieur. Plusieurs sciences pouvant être considérées comme étant de niveau supérieur, comme la biologie évolutionnaire ou la thermodynamique, sont d'ailleurs à bien des égards autonomes et elles n'ont pratiquement pas besoin de se référer à d'autres sciences ni à la constitution des systèmes à l'étude. Ainsi, ces systèmes présentent des niveaux d'organisation dont les entités, propriétés, états, etc., sont « robustes ».

Malgré cette autonomie explicative toute relative, il importe de s'interroger sur la relation que peut entretenir un ensemble d'entités au niveau inférieur, des « réalisateurs », avec celles au niveau supérieur, les « réalisés ». Par exemple, pourquoi plusieurs systèmes différents à un certain niveau peuvent exemplifier une certaine régularité à un autre, comme avec les phénomènes thermiques de systèmes composés de molécules très différentes, ou le cas de la douleur physique chez différentes personnes ?

Une façon de voir les choses est de dire que les détails permettant de distinguer des systèmes ou états comme étant différents sont largement « impertinents » (« *irrelevant* », Batterman 2002 : 73). Mais qu'ils soient pertinents ou non, ces détails sont parfois simplement inaccessibles car inobservables (section 2.2.3), comme c'est le cas avec les propriétés dynamiques des molécules en mécanique statistique (Chapitre 5). L'argument de Batterman (2000 et 2002) vise à montrer comment les propriétés des niveaux supérieurs, typiques des sciences spéciales, peuvent être « multiples réalisées ». Il identifie une méthode mathématique, largement utilisée en physique, montrant que, du point de vue de la théorie traitant d'entités au niveau inférieur, plusieurs détails microscopiques sont impertinents, c'est-à-dire qu'ils n'apportent rien à l'établissement de régularités phénoménologiques. En ce sens, la réalisabilité multiple, en définissant un ensemble d'états différents mais néanmoins compatibles avec certaines variables phénoménologiques caractérisant un système, permettrait d'expliquer le comportement de ce système de manière synchronique. Ainsi, cette stratégie explicative ne fait pas appel à ce que devrait être une loi, excluant par le fait même les ensembles disjonctifs, ou à ce que devrait être une réduction, excluant certaines asymétries. Elle fait plutôt intervenir une expectabilité (section 2.1.4) où il est « naturel » que plusieurs états (microscopiques) « mènent », au sens donné par la compatibilité, à un état donné (macroscopique).

Dans la conception interventionniste, on constate que la réalisabilité multiple implique que la vaste majorité des interventions au niveau microscopique ne font pas de différence au niveau macroscopique. De sorte que les interventions au niveau macroscopique sont « causalement plus efficaces » (Wimsatt 2012 : 79) à induire des changements à ce niveau que les interventions sur des variables à un niveau inférieur. En effet, typiquement, une intervention au niveau inférieur, microscopique, doit être « fine » et requiert donc des informations sur les propriétés de ce niveau. C'est peut-être précisément en raison de cette robustesse par rapport aux changements sous-jacents qu'il est permis d'établir une régularité causale et de parler d'un même effet pour une même cause. Autrement, d'infimes modifications de l'état d'un système entraînerait l'impossibilité d'identifier une condition suffisante à la causalité, ou plus simplement à des corrélations. Certains modes d'accès ne permettent pas une détermination complète des propriétés d'un certain niveau. Ainsi, la réalisabilité multiple peut être vue comme une conséquence de la définition même de différents niveaux d'organisation et de leur mode d'accès, et n'entraîne pas nécessairement l'échec d'une réduction, pour autant qu'elle ne soit pas définie comme une déductibilité terme à terme. Elle n'implique donc pas nécessairement l'émergence épistémologique et encore moins l'émergence ontologique (section 2.3).

2.3 Émergence

Dans son acception courante, le concept d'émergence fait souvent référence à une « apparition soudaine ». Cet aspect de soudaineté fait écho à son tour à ce qui paraît imprévu, inattendu, nouveau. Sans véritablement en donner une signification plus précise, ce concept se laisse aussi résumer par le slogan « le tout est supérieur à la somme des parties ». En quoi ce tout serait-il « supérieur » et dans quel sens serait-il « soudain » est bien entendu matière à réflexion et à débat. Mais un tout peut être « supérieur » à la somme de ses parties en ce sens où il ne peut se réduire à une simple agrégation ou addition de celles-ci, que sa caractérisation ne peut en conséquence être déduite ou prédite de celles des parties, bref si ses propriétés transcendent en quelque sorte celles de ses composants. De là à dire que la description des parties, avec leurs propriétés et les lois qui les régissent, est insuffisante à l'*explication* du tout il y a un pas qu'il semble naturel de franchir, mais qui demande plus de précisions. Quoiqu'il en soit, l'ensemble des concepts évoqués dans ce paragraphe – apparition, imprévisibilité, nouveauté, irréductibilité, non-aggrégativité, non-déductibilité, inexplicabilité – se retrouvent dans les débats philosophiques, comme celui du problème corps-esprit, visant à donner un réel sens au concept d'émergence.

Compte tenu du projet qui est le nôtre ici, soit celui de faire sens du concept d'entropie en ce qu'il se présente à *la fois* explanandum et explanans de l'émergence, l'accent sera mis sur l'analyse de l'articulation entre émergence et explication, c'est-à-dire de ce qui permet à l'émergence d'expliquer un phénomène et de ce qui permet d'expliquer un phénomène d'émergence, voire l'émergence en général. Tel que discuté précédemment, adosser ainsi la notion d'émergence à celle d'explication n'a rien de nouveau et remonte à la conception empirico-logiciste des explications réductives.

Le concept d'émergence a d'abord été élaboré, dans la littérature philosophique moderne, afin d'obtenir un compromis entre deux ontologies extrêmes mutuellement exclusives : une ontologie moniste et matérialiste, essentiellement réductionniste, où il n'y aurait que des éléments matériels et leurs propriétés, et une ontologie dualiste, donc antiréductionniste, où il y aurait la matière mais aussi l'esprit. Ainsi, bien qu'il n'y ait pas de consensus à son égard, le concept d'émergence est généralement reliée, de manière qu'il reste à préciser, au phénomène qui apparaît à partir d'un phénomène plus *élémentaire* tout en lui étant dans une certaine mesure *autonome*. Le concept philosophique d'émergence est donc un concept essentiellement *relationnel* : les entités émergentes apparaissent ou surviennent d'entités plus fondamentales mais sont tout de même nouvelles ou irréductibles ou inexplicables par

rapport à celles-ci²¹. Il s'agit donc de concilier une certaine relation de dépendance, un « enracinement » (« *goundedness* »), avec une relative indépendance, une certaine autonomie. La relation entre ce qui émerge et ce à partir de quoi cela émerge étant souvent explicitée comme une relation entre un tout et ses parties, bref en termes *méréologiques*²². Il y a donc plusieurs questions épistémologiques et métaphysiques reliées à la notion d'émergence.

D'abord, il y a la question même de l'*existence* de phénomènes vraiment émergents et non pas seulement en apparence, à l'instar d'épiphénomènes, c'est-à-dire de ce qui se surajoute à un phénomène sans exercer sur lui aucune réelle influence, autrement dit à des illusions. Par exemple, la beauté d'un tableau dépend de l'arrangement des molécules mais elle n'exerce aucune influence sur celles-ci ni ne semble être au même niveau de réalité, si tant est que la réalité puisse venir en divers degrés. En effet, il est souvent soutenu que l'une des conditions nécessaires à l'existence, au statut d'existant, est la possession de propriétés causales (et l'on suppose que ce qui est causal est réel). Un autre exemple est celui des propriétés dites *résultantes* (par ex. trois briques de 3 kg donnent un tout de 9 kg) sont seulement « calculées sur » les propriétés de bases (les trois briques), « découlant » simplement, arithmétiquement, de propriétés plus élémentaires.

Ensuite, on retrouve la question de sa *portée*, en ce sens où il peut s'agir d'une notion *ontologique* ou *épistémologique* (voire même purement *linguistique*). En effet, on peut estimer que certaines descriptions macroscopiques ne puissent être réduites à une description microscopique et en conclure, ou non, qu'elles correspondent, respectivement, à deux types d'entités existantes, ou encore qu'elles nécessitent seulement deux types d'explication, ou qu'il s'agit simplement de deux façons de parler de la même chose, etc. Mais ces questions sont intriquées, car il ne va pas de soi qu'une justification d'assomptions ontologiques soit possible sans recours à des assomptions épistémologiques. C'est pourquoi le concept d'émergence est souvent présenté comme un concept « négatif », c'est-à-dire comme l'échec ou l'impossibilité de satisfaire une relation épistémologique comme une explication ou une dérivation. Il y a aussi la question, reliée à la précédente, de la définition des *entités* émergentes, des *relata*. Du point de vue ontologique, ce peut être, de manière simplement illustrative, des propriétés, des substances, des processus ou des événements, ou des pouvoirs causaux, tandis que du point de vue épistémologique, ce peut être des prédicats, des concepts, des lois, des modèles ou des théories.

²¹ En caractérisant l'émergence par l'échec d'un certain type d'explication (comme ce qui permet de comprendre), on peut reformuler la citation de Miller (en exergue de ce chapitre) comme suit : « *Confusion* [emergence] is a word we have invented for an order which is not understood [explained] ».

²² Batterman (2000 et 2002) présente un contre-exemple notoire de l'émergence méréologique (discuté plus loin).

Enfin, bien entendu, la question, qui n'est pas étrangère à la précédente, de la *définition* même de cette *relation* qu'est censée représenter l'émergence, que ce soit en termes de nouveauté, par rapport à des pouvoirs causaux ou d'autres propriétés, d'irréductibilité, d'imprédictibilité, ou de tout autre critère pertinent. Une caractérisation partielle de cette relation, qui est aussi prise comme une condition nécessaire à l'émergence, est celle de la survenance, ou encore de la réalisabilité multiple (déjà discutées). Il s'agit d'une condition de dépendance méréologique où le tout est déterminé par les parties. Humphreys (1997) soutient plutôt que l'émergence est caractérisée par une « fusion » des parties formant le tout émergent où ses parties cessent d'exister en conséquence de cette fusion (discuté plus loin).

Il va donc de soi que les débats à propos de la notion d'émergence ont été nombreux et se poursuivent encore aujourd'hui (voir Silberstein 2002 ; Bedau & Humphreys 2008 ; Batterman 2009 ; Corradini & O'Connor 2010 ; O'Connor & Wong 2012). Bien que l'émergentisme ne soit pas foncièrement anti-réductionniste, il exprime néanmoins l'échec d'un certain réductionnisme, soit un réductionnisme moniste et fondationnaliste où il n'y aurait qu'un seul niveau réel et digne d'intérêt dans l'explication scientifique. L'entremêlement des notions de questions épistémologiques et ontologiques peut ainsi très bien mener à une confusion des genres. Faire sens de l'émergence exige donc des précautions et c'est pourquoi l'on se tourne vers une taxonomie de ce concept polysémique.

2.3.1 Taxonomie(s) de l'émergence

Devant la diversité des conceptualisations de l'émergence, il devient important de classer cette diversité. Humphreys (2008) offre une double taxonomie de l'émergence. La première est appelée la taxonomie *relationnelle* puisqu'elle concerne la relation entre l'entité émergente et les entités de laquelle elle émerge. Bien que ce soit souvent le cas, elle ne se limite toutefois pas aux conceptions méréologiques. Elle reprend cependant la distinction classique du problème de la connaissance, entre émergence *ontologique* et émergence *épistémologique*, cette dernière étant séparée entre émergence *inférentielle* et émergence *conceptuelle*. L'émergence inférentielle découle d'une impossibilité à prédire ou calculer l'état ou une certaine propriété d'un système sur la base d'une théorie concernant un « domaine »²³. L'émergence conceptuelle, en revanche, souvent reliée à la précédente selon l'auteur, découle de l'impossibilité de représenter adéquatement un état ou une propriété avec l'appareil

²³ Humphreys n'est pas très explicite toutefois à ce sujet. Un « domaine » peut être considéré comme l'ensemble des entités (en général concrètes pour une science empirique ou factuelle), de leurs propriétés et de leurs changements, dont traite une théorie (voir Mahner 2007).

conceptuel et descriptif d'un cadre théorique donné. La seconde taxonomie, qui recoupe la première, est la taxonomie *temporelle* puisqu'elle catégorise les phénomènes émergents selon qu'ils procèdent d'un processus dans le temps, soit l'émergence *diachronique*, ou encore de la coexistence simultanées d'entités, soit l'émergence *synchronique* (Rueger 2000 et Hüttemann 2005 reprennent cette distinction).

L'émergence est souvent conceptualisée en opposition au réductionnisme et c'est pourquoi la taxonomie de Silberstein (2002), qui se concentre sur la taxonomie relationnelle, est utile. Du côté ontologique, on retrouve l'émergence se définissant comme la *non-élimination* d'une entité ontologique, comme une propriété ou un type, mais aussi la *non-identité* d'une telle entité. Généralement, les définitions de la non-élimination et de la non-identité renvoient à des considérations épistémologiques, explicatives. Il y a ensuite les concepts opposés à la survenance, soit l'*émergence méréologique*, comme quoi les propriétés du tout ne sont pas déterminées par celles des parties, et l'*émergence nomologique*, où des entités ontologiques sont régies par des lois de niveau supérieur qui ne sont pas déterminées ou nécessitées par des lois au niveau inférieur ou plus fondamental. Il est important de souligner que dans le cas de l'émergence nomologique, il est impossible *en principe* de dériver ou de prédire les phénomènes de niveau supérieur à partir des phénomènes de niveau inférieur. Du côté épistémologique, il y d'abord l'*émergence explicative*, où les caractéristiques d'un tout ne peuvent, *en pratique*, être expliquées ou prédites par celles des parties, comme des lois régissant leur comportement – ce qui est à peu près équivalent à l'émergence inférentielle. Il s'agit donc d'une émergence *qua* échec de la stratégie R (vue précédemment). Même si un certain type d'explication réductive est possible et si, donc, il n'y a pas d'émergence explicative, il peut y avoir *émergence représentationnelle* ou *cognitive*, c'est-à-dire que certaines schèmes appliqués au niveau supérieur ou certaines régularités irréductibles à l'analyse des composantes ne peuvent être correctement représentés ou compris du point de vue du niveau inférieur. En somme, il y aurait un avantage pragmatique à conserver cette dualité épistémologique, cette discontinuité explicative. Pour autant, cela n'implique pas l'abandon d'un monisme causal ou d'une continuité déterministe (plus de détails plus loin).

Ces taxonomies peuvent être complémentées par une taxonomie des relata, soit ceux des *prédicats*, celui des *propriétés causales* et celui des *substances*. Schröder (1998) distingue plutôt les *choses*, les *propriétés* et les *lois*.

En situant ainsi le cadre de la discussion, on constate la diversité des approches philosophiques quant au concept d'émergence, mais il devient plus clair de saisir comment des théories de l'émergence en viennent à se positionner face à certains enjeux importants de la philosophie des sciences. Ces taxonomies ne servent qu'à apporter les distinctions nécessaires pour structurer la discussion. En effet, devant la diversité des approches et des propositions se rapportant l'émergence, il est facile de s'y

perdre et de se faire objecter d'être partiel ou partial. Autrement dit, il y a plusieurs thèses émergentistes différentes et bien entendu une conclusion philosophique sur l'une d'entre elles n'influence pas de la même façon une autre thèse émergentiste.

2.3.2 Conceptualisation de l'émergence

La discussion précédente a montré la diversité des approches du concept d'émergence. Celle qui suit vise à y offrir un cadre de conceptualisation à partir des principaux problèmes auxquels ces approches font face, dont celui de la caractérisation des relata (concepts, propriétés, processus ?), de la relation (méréologique ou non ?) et celui de la causalité descendante.

Dans un article célèbre, Kim (1999) présente ce qu'il considère les cinq thèses de l'émergentisme.

THÈSES DE L'ÉMERGENTISME

- a. *Émergence des entités complexes de niveau supérieur.* Les systèmes d'un niveau d'organisation supérieur émergent de la « rencontre » (« *coming together* ») d'entités de niveau inférieur dans de nouvelles configurations structurales, soit la nouvelle « parenté » (« *relatedness* ») de ces entités.
- b. *Émergence de propriétés de niveau supérieur.* Toutes les propriétés d'entités de niveau supérieur « découlent » (« *arise out* ») des propriétés et des relations qui caractérisent leurs constituants, certaines propriétés des systèmes complexes de niveau supérieur étant « émergentes » et d'autres, seulement « résultantes ».
- c. *Imprédictibilité des propriétés émergentes.* Les propriétés émergentes ne sont pas prédictibles d'une information exhaustive à propos de leurs « conditions basales » (« *basal conditions* »), tandis que les propriétés résultantes le sont.
- d. *Inexplicabilité/irréductibilité des propriétés émergentes.* Les propriétés émergentes, contrairement aux propriétés résultantes, ne sont ni explicables ni réductibles en termes de leurs « conditions basales ».
- e. *Efficacité causale des émergents.* Les propriétés émergentes ont des pouvoirs qui leur sont propres, en ce qu'ils sont irréductibles à ceux de leurs constituants basaux²⁴.

À la lumière de la taxonomie précédente, cette conceptualisation de Kim est un cas particulier de l'émergentisme. Bien qu'elles soient générales, toute théorie de l'émergence n'adopte pas

²⁴ Kim semble ici commettre un abus de langage, car peut-on vraiment dire que ce sont les propriétés, et non les entités, qui possèdent des pouvoirs causaux ?

nécessairement ces thèses, qui sont essentiellement méréologiques. Elles recourent d'abord, pour capturer le sens de nouveauté, les définitions ontologique et épistémologique de l'émergence. Or, la distinction entre l'émergence de type ontologique et l'émergence de type épistémologique est loin d'être transparente. Tel que mentionné, les justifications à la base de l'identification d'un cas d'émergence ontologique sont souvent de nature épistémologique. Évidemment, un cas de figure important et qui nous préoccupe particulièrement ici est celui de l'entropie. Sethna (2011 : 95) affirme ainsi que l'accroissement d'entropie est une « propriété émergente » car elle ne peut être *expliquée* par la mécanique classique, et par conséquent tous les faits de la nature ne seraient pas seulement « microscopiques » (pour une discussion détaillée, voir Chapitre 6). Faire sens de l'émergence par la notion d'explication exige d'abord de préciser le sens de cette notion, qui en reçoit plusieurs (section 2.1).

Huneman (2008) propose deux critères auxquelles doit satisfaire le concept d'émergence : (i) le *critère de l'adéquation empirique*, comme quoi il doit être en accord avec un ensemble de faits scientifiques, et (ii) le *critère de la non-trivialité*, c'est-à-dire que ce concept ne doit pas être instancié « trop » facilement, autrement tout ce qui existe ou presque tomberait sous ce concept. Comme l'adéquation empirique est définie par celle des faits scientifiques, la satisfaction du premier critère coule de source. Car, non seulement les relata, que ce soit les concepts ou les propriétés, sont définis par les théories scientifiques, mais le concept même d'émergence est souvent défini en fonction de certaines vertus épistémologiques comme l'explicabilité ou la prédictibilité. En revanche, le critère de la non-trivialité cherche à éviter la vacuité du concept. Un tel concept serait alors compatible avec des thèses auxquelles il est censé s'opposer, comme certaines formes de dualisme et de réductionnisme. Ces deux critères apparaissent donc tout à fait raisonnables et pertinents à l'encadrement de la conceptualisation de l'émergence.

2.3.2.1 RELATA

Examinons d'abord la question des relata de l'émergence. Les thèses de l'émergentisme présentées par Kim (1999) montrent bien la diversité des relata (voir aussi section 2.3.2.1). Étant donné le lien étroit qu'entretient la notion d'émergence avec celles de réduction et d'explication, il peut sembler mal avisé de restreindre l'émergentisme à un seul type de relata. Il faut tout de même les circonscrire afin d'éviter la vacuité du concept d'émergence. De plus, puisque les entités telles les processus ou les propriétés causales ne sont appréhendées que par les concepts nécessaires à leur représentation, il est naturel de s'interroger sur la pertinence de vraiment choisir entre des relata épistémologiques ou ontologiques.

Le cadre réaliste adopté plus tôt autorise les inférences de prémisses épistémologiques, comme celles à propos du concept d'atome et de son statut théorique et empirique, à des conclusions ontologiques,

comme l'entité réelle que serait l'atome. Mais cela ne signifie pas que toute inférence du genre soit légitime ou justifiée, ni même qu'elle soit désirable. Il a été avancé que les considérations d'ordre métaphysique pouvaient être utiles dans la systématisation de nos connaissances, non pas qu'elles étaient nécessaires.

Par ailleurs, il y a une relativité des relata dans le concept d'émergence méréologique en ce qu'une entité quelconque peut être la fois considérée comme « basique » et « émergente ». En effet, un niveau au sein d'une hiérarchie est aussi relatif, en ce sens qu'il peut être considéré à la fois « inférieur » et « supérieur » par rapport à d'autres niveaux. Par exemple, une cellule biologique peut être considérée comme une entité supérieure, potentiellement émergente, relativement aux molécules qui la composent et comme une entité inférieure, ou basique, par rapport à l'organisme vivant composé de telles cellules. Dans sa forme la plus simple, la hiérarchie se divise en deux niveaux souvent appelés *microscopique* et *macroscopique*. Une certaine forme de fondationnalisme est ainsi généralement adoptée, en ce sens où certaines composantes, généralement du ressort de la physique, constitueraient la base la *plus* élémentaire ou fondamentale. Cependant, la conception des niveaux d'organisation adoptée plus tôt n'exige pas une telle position.

Une question importante concernant les relata de l'émergence est celle entre les propriétés et les processus. Huneman (2008) par exemple, soutient que les propriétés sont émergentes seulement de manière corrélatrice, c'est-à-dire en tant qu'élément d'un processus émergent. (Un processus est défini comme une séquence chronologique d'états ; un état est un ensemble de propriétés à un instant donné ; donc un processus est une séquence d'ensembles de propriétés.) En ce sens, soumettre qu'un processus est émergent exige une condition plus forte que dans le cas d'une propriété émergente. Son argument en faveur des processus comme relata de l'émergence s'appuie sur l'exemple d'un automate cellulaire formé de cellules noires et de cellules blanches. Le motif (« *pattern* ») global formé de ces cellules n'est pas une propriété que possède les cellules mais est plutôt donnée par leurs relations ; par conséquent, il s'agirait d'une propriété émergente au sens de non-aggrégativité ; or des motifs peuvent être identiques même si les cellules individuelles sont dans des états différents, c'est-à-dire noires ou blanches. De sorte que des motifs identiques, considérés émergents selon une définition méréologique (ou « combinatoire »), seraient tout de même différents en ce que l'un des deux serait fixe et l'autre « clignotant », c'est-à-dire qu'il ferait partie d'un processus déterminé par une règle dont le résultat est une alternance de cellules noires et blanches à chaque étape itérative de la règle. Une définition méréologique de l'émergence, focalisant sur le motif, ne pourrait distinguer le motif fixe et celui « clignotant », mais, en tenant compte du processus, cette distinction pourrait être opérée.

Cependant, cet argument n'élimine pas la possibilité que les relata émergents puissent être *autant* des processus que des états ou propriétés. Il ne semble pas non plus y avoir de raison particulière à ce qu'une propriété considérée comme émergente d'un système caractérise son comportement dans le temps, bref le processus qu'il manifeste. L'émergence de cohérence, comme des oscillateurs harmoniques reliés entre eux et vibrant à la même fréquence ou une foule applaudissant de même, est un exemple, tout comme l'émergence de symétrie, d'un ordre structural, comme dans le cas du motif exprimé par l'ensemble des cellules de l'automate cellulaire. Il pourrait en être de même pour l'irréversibilité de certains systèmes thermodynamiques (Chapitre 6). Toutefois, il n'est pas clair si les propriétés d'un système composé manifestées par les relations entre ses composants, en bref les *propriétés structurales*, peuvent être des propriétés émergentes. Est-ce qu'un château de cartes est émergent ? La question n'est pas simple. Généralement, les propriétés structurales sont considérées survenir sur des propriétés intrinsèques et les propriétés émergentes sont parfois évaluées de la même façon. Alors, en conférant un statut différent, à préciser toutefois, aux propriétés structurales, il reste à déterminer les rôles épistémologiques qu'elles peuvent jouer, au sein d'une explication par exemple ou d'une relation d'émergence.

Ceci dit, accepter les propriétés structurales comme des propriétés émergentes apporte une solution et un problème. D'une part, cela permet de rendre compte de l'apparition de motif où les agents demeurent intacts ou inchangés, mais aussi de l'idée qu'une propriété émergente est en quelque sorte le résultat d'une « interaction essentielle » des constituants (notons par ailleurs que la notion de processus semble aussi en rendre compte de manière triviale). Mais, d'autre part, le simple fait que des agents s'arrangent ou s'agencent d'une certaine façon, même de manière éphémère, devrait pouvoir compter pour une propriété structurale. Afin de ne pas trop élargir la classe de phénomènes pouvant être comptés comme émergents (et de contrevenir au critère de non-trivialité), il faut donc un critère supplémentaire. Bien que cet aspect éphémère ou provisoire soit rarement discuté, un tel critère pourrait être défini autour de la notion de stabilité. Mais Rueger (2000) propose une définition de l'émergence exigeant plutôt l'instabilité structurale. Une autre façon serait de contraindre la manière, autrement dit le processus par lequel la propriété structurale est obtenue (Chapitre 6).

2.3.2.2 RELATION

On a dit qu'au sein de la taxonomie relationnelle, il y a relation entre l'entité émergente et les entités de laquelle elle émerge, lesquelles constituent ce qu'on appelle la « base » de l'émergence. La notion de base de l'émergence tire son origine de l'idée, sans doute intuitive, que les objets simples précèdent, ontologiquement ou conceptuellement, les objets complexes ou du moins composés d'objets plus simples. Mais dans l'entreprise de caractérisation de l'émergence, la description de l'entité émergente

est aussi l'aboutissement d'une inférence dont les prémisses sont formées d'une description de cette base. L'entité émergente, soit l'un des relata, émerge donc de quelque chose – la « base » – et elle doit entretenir avec celle-ci une relation appropriée pour être considérée comme émergente, car cette entité peut être instanciée sans pour autant être considérée comme émergente.

C'est pourquoi l'émergence est aussi un prédicat relationnel : il n'y a pas de propriétés émergentes dans l'absolu, pas d'émergence *ex nihilo*. Les deux types d'émergence temporelle présupposent aussi une certaine base d'émergence. Outre cet aspect de la taxonomie relationnelle, celle-ci peut être subdivisée selon deux aspects importants du problème philosophique de la connaissance, soit l'aspect ontologique, qui concerne les objets considérés réels, et l'aspect épistémologique, qui est relatif à nos modalités représentationnelles de ces objets. Au sein de la taxonomie relationnelle, l'émergence inférentielle et l'émergence conceptuelle concernent l'aspect épistémologique de l'émergence.

Le concept d'émergence est souvent dépeint par la notion de nouveauté permettant de capter cette notion d'autonomie relative. Bien qu'elle soit sans doute la propriété la plus commune, la nouveauté est néanmoins le plus « insaisissable » des critères d'émergence (Rueger 2000 : 301). Comme bien souvent les propriétés d'un système, d'un tout, sont nouvelles par rapport à celles de ses parties, le problème est de choisir *quelle* propriété est vraiment émergente. Il n'y a pas de critères clairs établissant les propriétés émergentes parmi l'ensemble potentiellement infini de propriétés nouvelles. Encore une fois, cette nouveauté est souvent caractérisée, le cas échéant, par des notions épistémologiques comme l'imprédictibilité ou l'explicabilité. Définir l'émergence par le critère de nouveauté sans plus de précision constitue un problème majeur puisqu'il ne permet pas de distinguer les propriétés réellement émergentes des propriétés résultantes. De sorte que tout système composé serait émergent, c'est-à-dire qu'il aurait une ou plusieurs propriétés émergentes, ce qui ne satisfait pas le critère de non-trivialité. Par exemple, sans plus de précision sur la signification du critère de nouveauté, on pourrait bien dire que la propriété « peser 9 kg » attribuée à un système composé serait « nouvelle » par rapport à la propriété « peser 3 kg » des constituants du système.

Bunge (2003 : 17) définit une propriété ontologiquement émergente d'une entité complexe comme une « propriété qualitativement nouvelle », c'est-à-dire une propriété que ne possède aucun de ses constituants ou précurseurs. Prenons l'exemple du cercle : un cercle est une collection de points et les points individuels n'ont pas de forme, de sorte qu'être un cercle serait la propriété du tout et non des parties. En ce sens, « être un cercle » serait une « propriété qualitativement nouvelle », donc une propriété émergente, contrairement, par exemple, à la propriété simplement résultante « peser 9 kg ». Mais Bedau (2008) soutient que, puisque la propriété « être un cercle » peut être *prédite* de l'information que tous les points sont équidistants d'un point donné, cette propriété serait une propriété résultante.

Selon ce dernier, la propriété « être un cercle », bien que résultante, ne serait que « nominalement émergente » dans un sens proche de celui de Bunge. La notion de nouveauté est donc insuffisante pour caractériser celle d'émergence et doit être complémentée par des attributs qui, plus souvent qu'autrement, sont de nature épistémologique.

La stratégie généralement employée afin de donner sens à la relation entre les relata d'émergence est d'abord de définir une hiérarchie de niveaux (section précédente) et par là même une relative discontinuité entre ces niveaux. Ce n'est pas unique à l'émergence car toutes les disciplines scientifiques, hormis peut-être la physique des particules, étudient des systèmes composés de plusieurs sous-systèmes – une conséquence des niveaux d'organisation définis plus tôt. Mais, au sein d'une conception de l'émergence, ces discontinuités ne peuvent être totales et « hermétiques », ou « étanches », au point de verser dans la négation de toute forme de continuité, et partant dans une position foncièrement antiréductionniste et dualiste. En revanche, il serait vain d'étioler ces mêmes discontinuités au point d'effacer toute hiérarchie et d'adopter une position franchement réductionniste et moniste. Car, dans les deux cas, cela reviendrait à se couper de l'étymologie même du concept d'émergence, mais aussi et surtout de l'idéal de médiation qu'il présuppose, le reléguant au statut de simple synonyme sans vertu épistémologique. Autrement dit, tel que mentionné, l'émergentisme se situe dans une position intermédiaire entre les extrêmes du réductionnisme et du dualisme. L'idée générale de l'émergentisme, du moins dans sa mouture méréologique, est alors d'associer à une entité émergente une base élémentaire de niveau inférieur relativement à la hiérarchie adoptée. Le défi consiste ainsi à dégager un espace conceptuel, d'une part, à cette relation ou influence entre niveaux malgré la discontinuité définissant l'émergence et, d'autre part, à cette indépendance ou autonomie relative de l'entité émergente malgré ce qui la lie à sa base plus élémentaire. En clair, il s'agit de concilier une relation de dépendance relative entre entités qui demeurent néanmoins relativement autonomes (Bedau 2008).

C'est pourquoi le concept d'émergence est compromis. D'un côté, il peut se définir en fonction d'une relation forte entre niveaux d'organisation, comme une influence causale, ce qui mène en définitive à contredire les lois de la physique, instanciées à un niveau plus fondamental. D'un autre côté, cette relation peut être occultée, mais alors la description de certains processus, de certaines régularités à un niveau supérieur apparaît redondante en n'offrant qu'une reformulation de processus microscopiques. Les efforts de conciliation de l'émergentisme ont ainsi souvent insisté sur une description « grossière » (« *coarse-grained* ») des entités émergentes sans « déranger » ou « violer » la dynamique ordinaire des niveaux plus élémentaires ou fondamentaux inscrits dans une description « fine » (« *fine-grained* »). Mais, à ce titre, il semble que la « force d'évolution » ou le « vrai travail causal » du monde se situe au niveau fondamental et que les descriptions qu'offrent les niveaux

supérieurs se limitent à des « résumés utiles » sans portée ontologique ou à des épiphénomènes sans vertu explicative. Et en se tournant vers une description « grossière » de certains niveaux supérieurs considérés comme émergents, se trouve exemplifié le concept de réalisabilité multiple (section 2.2.5), qui fournit un exemple de ce type de problèmes. En effet, tel que discuté, ce concept se définit pour un nombre beaucoup plus grand d'états possibles et distinguables à un certain niveau (micro-états) que d'états à un niveau supérieur associé (macro-états), principalement en raison de la plus grande précision du vocabulaire employé pour décrire les micro-états, et de l'identité numérique du système de niveau supérieur avec le système de niveau inférieur ainsi décrit. En retour, le niveau le plus élémentaire ou fondamental se voit imposer un fardeau épistémologique de taille car il doit présenter une « stabilité », une « résilience » ou encore une « robustesse » suffisante à la construction théorique et potentiellement à l'explication des niveaux supérieurs.

Il faut alors préciser au moins trois choses. D'abord, la conceptualisation des niveaux d'organisation proposée ici se démarque de l'approche fondationnaliste car elle insiste sur la robustesse des propriétés formant ces niveaux sans pour autant accorder un statut privilégié à leur position axiomatique.

Ensuite, il faut bien remarquer que les niveaux d'organisation fonctionnent rarement en vase clos, en ce sens où des entités caractéristiques d'un certain niveau peuvent intervenir à plusieurs autres niveaux dans une explication. Par exemple, un champ magnétique peut influencer la trajectoire d'une particule comme l'aiguille d'une boussole, une politique économique peut influencer des institutions comme des individus, et différentes interactions peuvent se produire entre une cellule, un organisme et son écosystème. Il s'agit de relations entre niveaux (on parle parfois de relations « trans-niveaux ») sans nécessairement que ne soit exemplifié le concept d'émergence. Partant, des références à l'environnement d'un système à l'étude, que ce soit un atome comme un individu, sont pratiquement inévitables et cet environnement est pour ainsi dire toujours défini comme une entité macroscopique, d'un niveau supérieur. Ces éléments épistémologiques militent *contre* des conceptions extrêmes comme l'atomisme ou le holisme.

Enfin, et voilà un point important, il arrive souvent dans les conceptions de l'émergence que la relation entre niveaux, ou entre les parties et le tout, ou plus généralement entre ce qui émerge et ce à partir de quoi cela émerge, demeure indéterminée en ce sens où l'émergence découle d'une impossibilité, c'est-à-dire qu'est spécifié ce que l'émergence n'est pas. L'émergence inférentielle, qui découle d'une impossibilité à prédire ou calculer l'état ou une certaine propriété d'un système (section 2.3.1), en est un exemple. Des cas particuliers de ce type d'émergence sont l'*émergence faible* (Bedau 1997 et 2008), où un état d'un processus computationnel est faiblement émergent si et seulement s'il n'y a pas de raccourcis menant à cet état, et similairement l'*émergence computationnelle* (Huneman & Humphreys

2008 ; Huneman 2008), où un processus (et corrélativement un état) est dit émergent s'il génère un « ordre » ou un « motif » de manière incompressible (« *PSPACE complete* »). Le critère d'incompressibilité est intéressant puisqu'il s'appuie sur une démonstration mathématique : il serait donc objectif puisqu'indépendant de nos capacités épistémiques.

Il y a toutefois quelques objections que l'on peut soulever contre cette conception émergentiste. Dans un premier temps, comme elle s'applique à des processus computationnel, à des règles formelles itératives, on peut se demander sur quoi s'appuie le choix de la règle. Quel intérêt y a-t-il à choisir une règle formelle, à première vue arbitrairement, et à vérifier si un certain critère d'incompressibilité s'y applique ? De sorte que l'instanciation de ce concept dans le monde n'est pas établie. Autrement dit, le monde fonctionne-t-il comme un ordinateur ?²⁵ Il semble en effet y avoir un ensemble beaucoup trop grand de règles donnant naissance à des processus émergents au sens computationnel du terme pour qu'il puisse satisfaire le critère de non-trivialité. Par exemple, un tel processus incompressible pourrait suivre une règle telle que « demander un nombre à chacune des personnes dans cette salle et écrire cette séquence de nombres ».

Dans un deuxième temps, ce qui est censé émerger, comme des motifs, sont aussi en nombre illimité. Dans l'exemple de l'automate cellulaire contenant des cellules noires et blanches, il se pourrait bien qu'après quelques itérations « émerge » une lettre chinoise dont la reconnaissabilité dépende des capacités épistémiques de l'observateur. Même en circonscrivant le domaine des processus possibles il semble impossible de limiter suffisamment, pour un système assez large, et même s'il est beaucoup plus petit qu'un système macroscopique à près de 10^{23} degrés de liberté, tous les motifs ordonnés possibles, surtout localement ; il paraît toujours possible, à titre d'exemple, qu'un motif constitué de quatre cellules noires formant un carré « émerge » quelque part. En voulant les limiter encore davantage il s'avérerait risqué d'éliminer certains processus émergents d'intérêt, comme l'apparition de la vie, localement, dans une « soupe chimique ». Même si ces exemples, le carré et la vie, sont très différents, la possibilité de les distinguer de manière apriorique risque fort de compromettre le modèle de la nouveauté entendu au sens d'incompressibilité.

Dans un troisième temps, il vaut la peine de rappeler ici que prédire n'est pas expliquer. Posséder une règle formelle incompressible menant à un certain processus ou à un certain état ne se résume pas à expliquer ce processus ou cet état. On peut arguer qu'une explication de type historique demeure possible mais elle paraît vaine si la règle est choisie arbitrairement. (De plus, ne *pas* prédire ne revient pas à ne *pas* expliquer.) Il est vrai toutefois que les efforts à la base de l'émergence computationnelle

²⁵ Voir Wolpert (2008) pour une discussion intéressante et rigoureuse.

visent surtout à donner un concept d'émergence et non pas à expliquer les phénomènes émergents. On peut néanmoins remarquer qu'une certaine explication est possible, même si elle paraît triviale : si une explication implique l'expectabilité, alors on peut dire que l'on est en droit à s'attendre à – donc d'« expliquer » – un phénomène nouveau et imprédictible – donc « émergent » –, car on est devant l'impossibilité de le prédire (de manière incompressible).

Un autre type d'explication possible est le suivant. Une simulation appelle bien sûr une règle itérative et des conditions initiales. Parfois des conditions initiales légèrement différentes peuvent mener à des résultats très différents. Lorsqu'un paramètre est modifié légèrement et que le résultat, comme l'émergence d'un motif, est très différent par rapport à celui sans cette modification même minime, alors on considère souvent que ce paramètre ou cette modification particulière est la *cause* de ce résultat²⁶. En supposant la modularité du système, cette approche peut être interprétée selon la stratégie M (section 2.1.1). En pathologie génétique, on identifie souvent une variante génétique, comme une mutation rare d'un gène, comme la cause d'une maladie (MacArthur *et al.* 2014).

En somme, l'émergence comme incompressibilité algorithmique propose un critère définitionnel objectif, et en cela il mérite l'attention, mais il ne permet pas d'identifier un motif ou une organisation d'intérêt, ni de le prédire, ni d'identifier clairement une règle menant à ce motif parmi une infinité de possibilités. C'est pourquoi cette proposition permet d'identifier clairement un processus *défini* comme émergent, mais en le décrivant sans l'expliquer (ou du moins, en un sens très faible ; section 2.1.4 et Chapitre 6).

2.3.2.3 CAUSALITÉ DESCENDANTE

La relation entre niveaux définissant cette dépendance relative peut être définie comme étant de nature causale, par ce qui est appelé la *causalité descendante* (« *downward causation* ») où les propriétés émergentes d'un système agissent de manière efficiente sur leur propre base d'émergence²⁷. Cette notion a toutefois été critiquée à plusieurs reprises (par ex. Kim 1999 ; Bitbol 2012). Elle a d'abord été taxée de *circularité vicieuse*, ou d'incohérence philosophique, car il est dit que des processus de niveau inférieur seraient responsables de processus de niveau supérieur, tandis que ces derniers exerceraient réciproquement une influence causale sur les premiers. Il semble en effet y avoir contradiction ou redondance dans l'affirmation à l'effet qu'une base d'émergence soit modifiée par ce qu'elle « supporte » ou « fait

²⁶ « *Try approaching simulations as experiments by changing only one variable at a time until you understand its effects on the outcome. Changing several things all at once can lead to confusion and an inability to see cause-and-effect relationships, unless you have fully understood the effects of individual variables.* » (Hamilton 2009 : 7)

²⁷ Dans une acception proche du sens aristotélicien selon Bitbol (2012).

naître ». Devant cette apparente contradiction, on peut distinguer la causalité descendante, sur le mode efficient et souvent métaphysique, de la contrainte descendante, sur le mode formel et généralement épistémique. Ensuite, si une entité à un niveau supérieur exerce une influence causale au niveau inférieur, comme celui de la physique, alors cette influence a pour conséquence de contredire la *fermeture nomologique* des lois fondamentales comme celles de la physique. Et une telle influence causale macroscopique devrait être incluse dans une explication des phénomènes microscopiques, mais si ces lois sont justement considérées fondamentales alors elle devrait mener à l'*exclusion* (explicative) de cette influence. Enfin, ce concept n'a reçu jusqu'à ce jour *aucun* support empirique.

L'approche interventionniste est adoptée par Bitbol (2012) pour interpréter l'émergence de même que les causalités descendante et ascendante afin de répondre, entre autres, à ce contre-argument de la circularité vicieuse. Il s'appuie sur ce qu'il considère être les leçons de la mécanique quantique et de l'idéalisme transcendantal kantien afin de se débarrasser de toute conception métaphysique et en particulier l'image dualiste et fondationnaliste. Ainsi, (i) une intervention à un niveau n serait *spécifique* et *contributive* à la *définition* de ce niveau et (ii) si une intervention à un niveau supérieur (n) est effectuée, certains phénomènes à un niveau inférieur ($n - 1, 2, \dots$) sont observés, tandis qu'un retrait de cet antécédent qu'est l'intervention à un niveau n implique que ces phénomènes à un niveau inférieur ($n - 1, 2, \dots$) ne sont plus observés. La causalité descendante est alors interprétée en fonction de modes d'intervention selon les termes causalistes de la stratégie M : une intervention sur X à un certain niveau d'organisation (n) cause Y qui est détectable par une autre intervention à un niveau d'organisation inférieur ($n - 1, 2, \dots$). Cette approche interventionniste permettrait ainsi, selon l'auteur, de solutionner certains problèmes associés à la notion d'émergence.

D'abord, puisque la question de la relation entre niveaux d'organisation est remplacée par celle de leur détectabilité, il n'y aurait pas de problème de circularité vicieuse associée à la causalité descendante. Ensuite, il n'y a pas le problème de l'influence d'un niveau supérieur sur un niveau inférieur, ce qui contredirait l'hypothèse de la fermeture nomologique des processus microscopiques, comme la mécanique et la dynamique. Enfin, la détermination d'une dualité de modes d'accès n'impose *pas* une *dualité ontologique* des niveaux d'organisation.

Il y a ainsi une différence majeure entre l'approche de Wimsatt (1981, 1994 et 2012) et celle de Bitbol (2012). Alors que le second n'infère pour ainsi dire rien d'une série d'interventions à un certain niveau, le premier s'autorise, d'après le critère de réalité susmentionné, à inférer la robustesse, ou son absence, de ce niveau et conséquemment la réalité de l'entité visée par l'intervention. Pour autant, il n'y a pas nécessairement de niveau fondamental présentant des propriétés intrinsèques et les deux approches peuvent très bien être considérées anti-fondationnalistes. Dans l'interprétation

interventionniste de l'émergence que présente Bitbol par contre, les niveaux d'organisation ne sont pas *réels* au sens *ontologique* du terme, mais sont plutôt *objectifs* pris dans son acception *transcendantale*, c'est-à-dire relatifs à un mode d'accès. Or, contrairement à la conception de Wimsatt, les modes d'accès sont essentiellement empiriques, au sens premier d'intervention, et évacue ce qui est inobservable. En ce sens, ils ne fournissent pas de base d'inférence ni à la réalité des entités qu'ils visent, ni à la description d'entités autres, comme, justement, des inobservables. En somme, les interventions constituent ici une source d'information et de justification privilégiée, une base d'inférence légitime ou épistémologiquement fondée.

Il y a lieu de douter cependant que l'approche interventionniste solutionne vraiment les problèmes de l'émergentisme plutôt que de les éluder. D'abord, si une intervention à un certain niveau serait spécifique et contributive à la définition de ce niveau, il n'y a cependant rien qui permette de circonscrire les interventions, les modes d'accès, spécifiques à un niveau donné, de sorte qu'il y a, en principe, autant de niveaux que d'interventions. On se retrouve donc devant un pluralisme exacerbé des niveaux d'organisation. À la limite, des interventions comme des mesures de température et de pression d'un gaz ne pourraient être unifiées au sein d'une description ciblant un seul et même gaz. Autrement dit, par la loi de Leibniz à l'effet que les choses ne partageant pas les mêmes propriétés sont différentes, les niveaux d'organisation sont aussi différents que le sont les propriétés attribuées au moyen des interventions. Ce que se refuse au final cette approche est une unification théorique des résultats des interventions au sein d'un schème prédictif.

Ensuite, en définissant les interventions comme une source d'information et de justification privilégiée, pour ne pas dire la seule qui le soit, on est en droit de se demander si l'on ne confond pas justification et corroboration empirique, d'une part, et conceptualisation et heuristique, d'autre part. Il est en effet possible de justifier une croyance même si celle-ci n'a pas l'aval de l'expérience, interventionniste ou non. Et si l'on convient que l'expérience et plus spécifiquement l'expérimentation par des interventions constituent des bases d'inférence importantes pour la formation de concept, elles ne sont pas les seules. Le cas de la mécanique statistique est à cet égard probant (Chapitre 5). Les inférences statistiques sur les grands ensembles de molécules ne découlent pas du tout d'interventions, mais au contraire sont utilisés pour expliquer les résultats expérimentaux qui impliquent généralement des interventions. Il est en effet difficile d'imaginer une intervention sur des entités exemplifiant la réalisabilité multiple.

Même si la conception interventionniste de la causalité possède des mérites évidents, il est tout à fait possible que la causalité ne soit pas le concept adéquat pour témoigner du lien de dépendance entre niveaux d'organisation. Comme le font remarquer O'Connor & Wong (2005), le problème de la

causalité descendante se dissout si l'on considère l'émergence non pas comme une relation synchronique entre un tout et ses parties (une unité et ses sous-unités), mais plutôt comme un processus dynamique, une relation diachronique. Ainsi, il est préférable d'interpréter la causalité descendante sur le mode formelle où le niveau supérieur agit à la manière d'une intervention, d'une perturbation, au sens où il y a variation d'une condition limite, délimitant le domaine de variation des sous-unités ou sous-systèmes sans toutefois dicter cette dynamique.

2.3.3 Conclusion préliminaire

Le concept d'émergence peut être défini comme l'occurrence de propriétés qualitativement nouvelles, mais cette définition doit être précisée et doit aussi satisfaire les critères de non-trivialité et d'adéquation empirique. Ce faisant, le prédicat d'émergence peut s'appliquer à des états ou à des processus à différents niveaux d'organisation (section 2.2). Comme un niveau d'organisation est une description d'un ensemble de propriétés et de régularités suffisamment stable pour être déterminé de manière robuste, selon divers modes d'accès, l'application du prédicat d'émergence n'implique donc pas un fondationnalisme ni une réification de ces niveaux, donc nul besoin de l'harnacher à une réduction totale. C'est pourquoi la définition de ces niveaux découle d'une systématisation des connaissances obtenues par ces différents modes d'accès, autrement dit des théories scientifiques qui sont censées refléter la structuration effective des processus naturels. En même temps, cette structuration conditionne largement la façon dont sont systématisées nos connaissances, car il est effectivement possible d'exprimer des phénomènes apparemment divers avec des descriptions générales et des lois relativement simples. En somme, la cadre réaliste adopté ici ne fait que soutenir la thèse à l'effet qu'il y a une raison (la structuration effective de la réalité ; mais *laquelle* ?) pour laquelle la réalité est connaissable et que nos connaissances, pourtant intrinsèquement faillibles, offrent la possibilité, sans garantie néanmoins, d'une description vraie de la réalité.

C'est pourquoi l'émergence peut être de type épistémologique ou de type ontologique. Le statut de cette dernière étant tirée d'une inférence de la première. Malgré les difficultés intrinsèques de soutenir une telle inférence, elle n'est pas récusée d'emblée. L'émergence ontologique est toutefois aux prises avec deux problèmes de taille. L'un de ces problèmes est de distinguer les propriétés émergentes des propriétés ordinaires, comme avec la notion de nouveauté, et pour ce faire il est, semble-t-il, impossible de faire l'économie des théories scientifiques. L'autre problème est de spécifier la relation entre niveaux, car si l'émergence correspond à la structuration effective des processus naturels, elle doit

expliciter ce qui relie les propriétés basiques et les propriétés émergentes. Encore une fois, cette relation, causale ou autre, correspond généralement à la relation entre explanans et explanandum.

L'émergence peut donc aussi être considérée comme un principe *méthodologique* suggérant la recherche d'explication réductionniste autant que possible, sans multiplier les entités théoriques et sans pour autant coucher les phénomènes sur un lit de mystère, tombant par exemple dans un pluralisme exacerbé des niveaux d'organisation selon leurs différents modes d'accès. Affirmer qu'un phénomène est émergent, c'est affirmer qu'il existe un lien modal (par ex. la survenance) entre des entités à des niveaux d'organisation différents sans qu'une explication réductive ne soit possible (par ex. il n'y pas de température sans atome). En ce sens, identifier un cas d'émergence revient à caractériser (expliquer ?) une situation théorique particulière où différents modes d'accès permettent de définir des niveaux d'organisation sans qu'il n'y ait de réduction totale (identité ou élimination). Ceci dit, comme il n'y a pas de moyen suffisamment consensuel de juger de la valeur des modes d'accès (est-il préférable de calculer ou d'observer ?), la possibilité d'une explication réductive n'élimine pas totalement la légitimité d'une affirmation niant le succès de la réduction pour des raisons cognitives ou représentationnelles ou pragmatiques, comme quoi, par exemple, il serait plus intuitif et sans doute plus simple de parler de chaleur plutôt que d'énergie cinétique des molécules.

2.4 Explication et émergence

Le concept d'émergence peut donc être caractérisé de plusieurs façons. Rappelons qu'une distinction importante est celle entre émergence épistémologique, soit en gros l'imprévisibilité des niveaux supérieurs, et émergence ontologique, comme l'occurrence d'une nouveauté qualitative. Encore une fois, la caractérisation de cette nouveauté s'appuie généralement sur des considérations épistémologiques revenant aux notions d'imprévisibilité ou d'explicabilité. Ce concept d'émergence étant, dans la plupart des conceptions, relationnel, il a été vu aussi que les relata peuvent être très divers (propriétés, concepts, etc.) et que donc la caractérisation des termes de la distinction précédente n'est pas unique. Mais peu importe l'approche adoptée, le concept d'émergence résultant et les différents éléments participant à sa définition seront nécessairement appliqués *au meilleur de nos connaissances*. C'est pourquoi l'approche épistémologique adoptée, la réponse au problème général de la connaissance, conditionnera en grande partie la conception de l'émergence retenue.

Que nos connaissances soient insuffisantes à inférer des assomptions ontologiques n'implique pas que l'émergence ontologique soit inexistante. Mais, encore une fois, cette dernière est largement influencée par nos connaissances scientifiques, entre autres éléments importants pour la définition de ses relata, des niveaux d'organisation (de réalité). En revanche, un cas d'émergence ontologique, définie comme la réalisation d'une « nouveauté » dont les termes doivent être précisés, peut bien impliquer l'impossibilité d'une relation épistémologique comme une réduction ou une explication réductive. Mais il est aussi possible de détacher le concept d'émergence de la notion d'explication, de définir cette nouveauté sans cette dernière, où un phénomène émergent pourrait tout de même recevoir une explication. Ceci dit, en définissant l'émergence ontologique comme l'occurrence d'une propriété nouvelle, sans plus, il y a lieu de douter que ce soit l'émergence qui explique le phénomène émergent. Mais il y a plus. Si l'émergence ontologique est définie comme la réalisation d'une nouveauté qualitative objective, même si elle implique l'impossibilité d'une explication réductive, elle n'implique pas nécessairement l'impossibilité d'une explication dans un sens plus large. Par exemple, la chimie quantique présente les mécanismes de formation, donc une explication mécaniste, du dihydrogène à partir de l'hydrogène monoatomique, alors que le premier présente des propriétés différentes, que l'on pourrait qualifier d'« émergentes », par rapport au second. En ce sens, « l'émergence expliquée est encore de l'émergence » (Bunge 2003 : 21).

Il faut alors faire preuve de prudence afin de ne pas opérer un glissement de sens d'un cas d'émergence expliquée à celui d'une émergence qui explique. Dans un premier temps, il faut distinguer si ce qui fait l'objet de l'explication, l'explanandum, est un phénomène émergent particulier ou plutôt

l'émergence en général, comme s'il s'agissait d'un processus naturel. Cela fait, pour que l'explication opère, il va de soi que l'émergence ne peut être définie en termes d'inexplicabilité, comme l'émergence explicative (présentée plus tôt). Il n'est pas clair cependant si une telle définition de l'émergence, décrivant une situation où il est impossible (en pratique ou en principe) d'expliquer certains types de phénomènes, implique l'impossibilité d'expliquer l'émergence elle-même. Dans un second temps, pour que l'émergence puisse expliquer il faut pouvoir l'inclure au sein d'une stratégie explicative, et cette stratégie reste à déterminer devant le pluralisme (restreint) décrit plus tôt. La question reste ouverte à savoir si l'émergence peut jouer un rôle essentiel au sein d'un explanans au même titre qu'une loi ou une cause. Ce qui est plus susceptible de se produire est une prétention à l'explication d'un phénomène donné en le subsumant sous un concept d'émergence. Le résultat est décevant puisqu'il revient à stipuler plutôt qu'à expliquer²⁸.

Devant la très grande diversité des conceptions de l'émergence et de l'explication, une certaine conception de l'émergence articulée autour de la notion d'explication, un modèle, sera utilisé afin d'étayer la prochaine discussion – un « banc d'essai » en quelque sorte. Ce modèle, désigné par la « stratégie R » (pour réductionniste) et qui a été proposé par Hüttemann (2005), est somme toute assez répandu. Selon lui, l'émergence découle de l'impossibilité d'une explication déductivo-nomologique de type réductif, et ce type d'émergence impliquerait une émergence ontologique, une sorte d'émergence nomologique.

2.4.1 Un modèle : la « stratégie R »

Hüttemann (2005) propose le concept de « *microexplication* » comme un type particulier d'explication déductivo-nomologique, soit une explication *méréologique* et *réductive*, devant satisfaire certains critères d'ordre général où certaines « informations » sur le tout et ses parties doivent être disponibles : d'abord, l'information sur la manière dont les parties se comporteraient si elles étaient isolées, ce qu'il appelle les « lois d'évolution », ensuite, l'information sur la façon dont ce comportement contribue à celui du système composé, soit les « lois de composition », et enfin, l'information sur les types d'interaction qui apparaissent si les parties ne sont plus isolées, les « lois d'interaction ». La stratégie explicative de la dynamique d'un système composé, l'explanandum, de Hüttemann (2005) est donc donnée par :

²⁸ Tel que mentionné, je vais soutenir un peu plus loin qu'il peut y avoir explication au sens faible.

STRATÉGIE R. L'état d'un système composé (*explanandum*) est expliqué s'il peut être décrit par (i) des lois d'évolution de ses parties isolées, (ii) des lois d'interaction et (iii) des lois de composition du tout par les parties (*explanans*).

Il distingue deux types de microexplication, qui consiste à expliquer le comportement du système composé, le tout, par celui de ses constituants, les parties.

Une *microexplication synchronique* est ce qui explique pourquoi un système composé est dans un certain état à un temps t en fonction des états de ses N constituants au temps t . Il s'agit d'une fonction $MS : E^{1...N} \rightarrow E^*$, où l'état E^* , caractérisé par une certaine quantité comme l'énergie par exemple, du système macroscopique, soit l'*explanandum*, est expliqué en fonction des états $E^{1...N}$ des systèmes microscopiques et des lois de composition, ce qui constitue l'*explanans*. C'est pourquoi ce type d'explication peut être reconstruit selon le modèle déductivo-nomologique.

Une *microexplication diachronique* est ce qui explique pourquoi un système composé, un tout, est dans un certain état à un temps t' en fonction de l'état précédent t du système et de sa dynamique, laquelle est liée à la dynamique des parties. L'*explanandum* correspond à la dynamique du tout et l'*explanans* correspond à celle des parties. Il s'agit ainsi d'une fonction telle que $MD : E^{1...N} \times T \rightarrow E^* \times T$. Dans une microexplication diachronique, il n'est pas nécessaire que les *états* des constituants (parties) soient spécifiables, ce qui importe ce sont les lois.

Le cas d'espèce pour Hüttemann (2005) est l'intrication quantique (voir aussi Esfeld 2004 ; Jodoin 2010a ; Bitbol 2012 ; Chapitre 5). L'état d'un système quantique est représenté par un vecteur d'état qui est un vecteur dans un espace mathématique dit de Hilbert dont les axes représentent les résultats de mesure possibles pour ce système, dans des circonstances données. La description généralement acceptée d'un système composé dit « intriqué » stipule que le vecteur d'état de ce système composé ne peut pas être factorisé en un vecteur d'état consistant en une somme des vecteurs d'états individuels composant ce système. En conséquence, l'état du système composé ne peut être représenté par le produit tensoriel des espaces de Hilbert associés aux systèmes individuels mais plutôt par leur superposition. De sorte que les probabilités d'un résultat de mesure sur l'un des systèmes composants, lesquelles sont exprimées par leur vecteur d'état individuel, dépendent des probabilités des autres systèmes individuels, même s'il n'y a pas d'interaction entre eux. Alors, si le vecteur d'état du système composé représente bien l'état de ce système comme un tout, en ce sens où il définit ses propriétés, alors il est tout fait raisonnable de dire que ce genre de système quantique est un exemple où les propriétés du tout qu'est le système composé ne se réduisent pas aux propriétés intrinsèques de ses composantes les plus élémentaires, en l'occurrence aux vecteurs d'état des composants. Selon la taxonomie précédente il s'agit d'un échec de la survenance méréologique. Tel que mentionné, l'échec

du réductionnisme, ou plus précisément d'un certain type de réductionnisme, est généralement considéré comme un cas d'émergence, ou plus précisément d'un certain type d'émergence. Il s'agit en somme de la proposition de Kronz et Tiehen (2002).

Mais ce n'est pas ce que propose Hüttemann. Plutôt, l'émergence découlerait de l'échec de la stratégie R, qui peut prendre deux formes. L'auteur soutient que si les cas d'intrication quantique mettent en échec la microexplication synchronique, ils n'affectent en rien la microexplication diachronique. L'intrication quantique répond ainsi aux critères d'une microexplication particulière et la mécanique quantique serait aussi réductionniste, en ce qui concerne les explications dynamiques, que sa contrepartie classique. Par conséquent, selon cet argument, l'échec d'une microexplication de type synchronique (diachronique) n'implique pas l'échec d'une microexplication de type diachronique (synchronique). Mais il fait une inférence de plus qui montre à quel point sont intriqués (justement) les aspects épistémologique et ontologique du réductionnisme et de l'émergentisme : selon Hüttemann (2005 : 115), l'émergence est une « notion *ontologique* » de type méréologique, qui se définit par l'échec d'une microexplication, synchronique ou diachronique, donc de la stratégie R. Il y aurait donc une émergence synchronique et une émergence diachronique.

D'abord, ce modèle est bien entendu limité par le modèle déductivo-nomologique de l'explication. Ainsi, on peut bien envisager d'autres explications de type réductif que les microexplications synchronique et diachronique de la stratégie R. Ensuite, l'inférence de l'émergence épistémologique de type explicatif à l'émergence ontologique requiert une justification, qui fait défaut ici. Ce type d'inférence demande une très grande confiance dans les théories scientifiques censées rendre ce type d'explication, en ce sens où l'impossibilité d'une théorie, en l'occurrence la mécanique quantique, à expliquer un phénomène consiste en une impossibilité de principe.

2.4.2 Émergence expliquée

Il y a deux sens à l'expression « émergence expliquée ». Il peut s'agir, d'une part, d'un phénomène exemplifiant un certain concept d'émergence, épistémologique ou ontologique, lequel recevrait une certaine explication. Ou encore, il peut s'agir, d'autre part, de l'émergence elle-même qu'il faudrait expliquer, en d'autres termes de dire pourquoi il y a des phénomènes émergents en général. Dans le premier cas, deux exigences encadrent cette possibilité : d'abord, l'émergence doit, trivialement, recevoir une conceptualisation, une définition allant au-delà de l'échec d'une explication et, ensuite, recevoir une explication de l'explanandum représentant ce cas d'émergence selon la définition préalablement adoptée. Autrement dit, par définition l'émergence explicative (section 2.3.1) ne peut

être expliquée et ne constitue en somme qu'un « signe de notre ignorance » (Bedau & Humphreys 2008 : 6). C'est le cas notamment de l'échec de la stratégie R. En distinguant deux types d'explication réductive (ou microexplication), soit le type synchronique et le type diachronique, on peut être porté à prétendre à une plus grande exhaustivité dans les possibilités explicatives, voire même avancer que ces microexplications épuisent les possibilités d'une explication réductive. Si tel est bien le cas, alors une situation d'échec de microexplications, donc de la stratégie R, est un cas d'émergence explicative. Autrement, pour expliquer un tel cas d'émergence, il faudrait donc une explication qui ne soit pas une explication de type réductif, qui ne soit pas une microexplication.

L'exemple de l'intrication quantique, utilisée pour présenter la stratégie R, est donc intéressant puisque Hüttemann (2005) avance qu'il est impossible *en principe* d'expliquer, au sens nomique du terme, le comportement du système composé par celui des parties composantes. Cette impossibilité de principe découle toutefois de la théorie au sein de laquelle sont décrits ce système composé, ses parties et les lois qui les régissent. En ce sens, cette émergence n'est pas différente de celle qui se présente au sein d'autres théories scientifiques, comme quoi l'irréversibilité thermodynamique serait « émergente » par rapport à la mécanique classique, et la vie par rapport à la physique voire même par rapport à la biologie. On peut poser que la mécanique quantique place un point final aux explications du domaine microscopique et qu'aucune autre théorie ne pourra jamais offrir d'explication réductive, de microexplication.

Mais il y a sans doute une raison plus fondamentale, ou simplement de nature différente, à cette affirmation d'une impossibilité de principe au-delà de la confiance que l'on porte en cette théorie : que la non-localité, à la base de cette intrication quantique, est propre au niveau quantique et complètement étrangère à « notre » niveau ou radicalement incompatible avec une certaine conception du monde, une conception mécaniste au sens classique du terme. C'est-à-dire que cette non-localité est tellement en porte-à-faux de nos conceptions courantes sur le monde qu'en rendre compte exige une rupture conceptuelle, voire métaphysique, qui ne peut que mener à ce qui été appelé l'émergence représentationnelle. Ainsi, même si une preuve de l'impossibilité de principe d'une explication réductive est difficile à obtenir, comme dans le cas de l'incompressibilité avec l'émergence computationnelle, il reste que des raisons cognitives ou représentationnelles laissent planer un doute non pas sur l'explication en général, mais sur l'explication réductive en fonction de concepts hétérogènes ou plus difficiles à comprendre ou simplement parce qu'ils semblent plus adaptés à un certain niveau d'organisation. Par exemple, désigner une tasse à café par l'arrangement de ses molécules est loin d'être pratique. Dans pareils cas, et tel que discuté (section 2.1.3), le constat d'échec d'une réduction procéderait de considérations pragmatiques. Par conséquent, l'émergence serait inférée à partir de cet échec même si une explication est disponible.

Malgré ces difficultés, un cas d'émergence peut aussi être expliqué si l'on détache la définition de l'émergence de la notion d'explication, donc si l'on évite l'émergence explicative. C'est le cas notamment de la proposition de Bunge (2003) définissant une émergence méréologique, dont les relata sont les propriétés, en termes de « nouveauté ». L'exemple précédent de la chimie quantique refléterait ce type d'émergence. Comme l'émergence est définie dans ce cas-ci par le critère de nouveauté, qu'il faut préciser toutefois, un cas d'émergence peut être expliqué par une explication de type réductif ou tout autre type d'explication jugée pertinente. Car, la notion de réduction n'est pas, *prima facie*, incompatible avec celle de nouveauté.

Pour O'Connor & Wong (2005), si une propriété émergente doit être expliquée, elle doit l'être de manière causale, et non de manière simplement formelle. Mais faire appel ainsi à la causalité engage tout un lot de problèmes intrinsèques à cette notion dont les auteurs ne traitent pas. Puisque l'explanandum est la propriété émergente, celle-ci doit être l'effet et la cause la base émergente. Il s'agit donc d'un cas de causalité ascendante, différent mais tout de même connexe à celui de la causalité descendante (section 2.3.2.3). D'abord, la causalité est une notion asymétrique et généralement chronologique, impliquant une succession invariable dans le temps, où la cause précède l'effet, mais comme les niveaux inférieurs et supérieurs sont contemporains, on voit mal comment les articuler au sein d'une relation causale. L'acausalité de la réalisabilité multiple, typique des systèmes complexes, repose aussi sur cet aspect synchronique des niveaux d'organisation. Mais à cela on peut ajouter qu'un ensemble disjonctif peut difficilement former un lien causal entendu comme une correspondance un-à-un de la cause et de l'effet. Au-delà des difficultés à définir la causalité en général s'ajoutent donc celles des niveaux d'organisation. Ensuite, ces auteurs ne précisent pas pourquoi l'explication devrait être de type causal. La principale raison pour laquelle l'explication ne peut être seulement formelle est que l'émergence qu'ils défendent est ontologique, en ce qu'elle dépend d'une disposition physique des parties à générer certaines propriétés émergentes d'un tout. Mais la motivation à la base d'une telle conception émergentiste se positionnant par rapport à la notion d'explication est beaucoup plus subtile. Il faut revenir à la motivation initiale derrière le concept d'émergence, se voulant comme un compromis entre deux ontologies extrêmes mutuellement exclusives, l'une moniste et l'autre dualiste.

Une position moniste quant aux substances ou aux propriétés tend à ramener la diversité des phénomènes à une classe unifiée et homogène de substances ou de propriétés, comme celles de la physique dans le cas du physicalisme (à ne pas confondre avec le matérialisme). C'est en fait le sens premier que l'on accorde à la notion de réduction où est abandonné l'affirmation d'existence ou de réalité de certaines entités, comme dans le cas de la thèse physicaliste à propos du domaine mental. Il est donc tout fait raisonnable de croire que l'on puisse expliquer cette diversité à partir d'entités et de lois fondamentales comme les atomes et les lois de la physique. Pour une position dualiste, en revanche,

cette possibilité est compromise car les entités à des niveaux différents respectifs sont par définition exclusives et hétérogènes (à tout le moins non identiques). Dans les deux cas l'émergence (méréologique et synchronique) est exclue : dans le premier, il n'y a qu'un seul vrai niveau, que l'on considère réel, et dans le second, il n'y a pas de lien modal, cet « enracinement » entre niveaux d'organisation. Ce lien modal est justement offert par une relation causale, mais il a été vu (section 2.3.2.3) que la causalité entre niveaux d'organisation est très problématique.

Une cause est censée offrir les conditions nécessaires et suffisantes à son effet et, en ce sens, elle soutient l'expectabilité recherchée dans une explication. Donc un phénomène émergent, et non l'émergence elle-même, serait expliqué en présentant sa cause provenant de la base d'émergence. Cependant, en plus des problèmes associés aux niveaux d'organisation synchroniques, dans le cas de l'intrication quantique (discuté plus tôt), si l'émergence peut recevoir une explication causale, les difficultés rencontrées sont celles associées à une description séparée des parties et du tout. Sans entrer dans les détails théoriques (voir les références à la section 2.4.1), lorsque deux systèmes quantiques ont interagi dans le passé, il n'est généralement plus possible de factoriser le vecteur d'état de la paire qu'ils forment (tout) en deux vecteurs d'état correspondant à chacun d'entre eux (parties) puisque le vecteur d'état du tout ne peut être formé d'un produit tensoriel. De sorte qu'une intervention sur une partie modifie les autres parties *et* le tout. Par conséquent, le concept de causalité entre niveaux d'organisation semble « insensé dans le domaine hautement holistique décrit par la mécanique quantique » (Bitbol 2012 : 244 ; voir aussi Jodoin 2010d). Donc, la stratégie M ne peut y être appliquée pour expliquer l'émergence des propriétés des niveaux supérieurs. Il peut paraître tentant par contre de se dédouaner de cette conséquence en affirmant que la mécanique quantique demeure, intuitivement et philosophiquement, un « mystère » (Gibbins 1989 ; Bricmont 2009). Mais, comme le souligne van Fraassen (1991 : 372), dire d'une chose qu'elle est un mystère est moins une assertion qu'une demande, une « demande pour une explication ». Ce qui se dégage donc de cette analyse est, tel que mentionné, la difficulté d'offrir non seulement une explication réductive mais aussi une description d'une relation de dépendance dans un cas de concepts aussi hétérogènes, ce qui peut se résumer à un cas d'émergence représentationnelle.

Le cas de la version statistique du concept d'entropie est aussi intéressant à l'analyse des relations entre niveaux d'organisation (Chapitre 5). Car cette version ne présente pas, *prima facie*, une propriété pourtant jugée essentielle au concept d'entropie thermodynamique qu'elle devrait représenter, voire même expliquer. Alors que ce dernier exemplifie une irréversibilité (Chapitre 4), le premier est incapable de le faire (Chapitre 5). C'est-à-dire que l'appareillage théorique utilisé pour définir le concept statistique de l'entropie ne semble pas posséder les outils nécessaires pour rendre compte de cette irréversibilité qui, dans le cas de la théorie réductive, revient à définir une fonction dynamique

non-décroissante de l'espace des phases. Pour y arriver, il faut faire appel à des concepts « externes », orthogonaux à cet appareillage théorique qu'est la mécanique classique. Ainsi, on peut faire appel bien sûr aux probabilités, mais des hypothèses supplémentaires, comme celle définissant un état équiprobable ou encore « typique », posent problème. D'autres hypothèses ne semblent guère plus justifiées, comme la méthodologie du grenage grossier ou la théorie ergodique. Comme on semble incapable de reproduire les propriétés désirées et que l'« image » du concept est obtenue au moyen d'hypothèses contestables, et qu'en plus y soit exemplifié la réalisabilité multiple, il semble bien qu'il s'agisse ici d'un échec réductionniste (Callender 1999). D'où l'affirmation que l'entropie est une propriété émergente (comme le montre le commentaire de Sethna 2011, présenté plus tôt ; voir Chapitre 6).

Une problématique supplémentaire dans l'entreprise d'explication d'un phénomène d'émergence est celle de bien cerner l'explanandum, soit la propriété émergente. Car s'il s'agit d'identifier une « discontinuité conceptuelle » à la base de l'échec réductionniste, comme une nouveauté qualitative, celle-ci peut se situer au niveau supérieur, c'est-à-dire au sein de la théorie macroscopique (par ex. la thermodynamique), ou plutôt dans ce qui sert d'intermédiaire entre différents niveaux de description, dans les lois-ponts ou dans cette « image » du concept que l'on tente de reproduire, ou même dans l'interprétation de cette « image ». Encore une fois, l'exemple de la mécanique statistique est révélateur. Il a été dit que le concept statistique de l'entropie fait appel aux probabilités (Chapitre 5), et comme les probabilités sont orthogonales ou inédites pour la mécanique classique, on peut penser que la discontinuité recherchée est celle de ce même concept, de sorte que le niveau émergent soit celui du caractère probabiliste des grands systèmes. On pourrait alors dire que ce dernier est mésoscopique et le niveau thermodynamique macroscopique. Mais, pour différentes raisons (que je laisse de côté pour l'instant), il semble parfois nécessaire d'interpréter ce concept réductif et c'est le cas du concept statistique de l'entropie, auquel on accole des notions de « désordre » ou « d'information manquante » (Chapitre 5). La propriété émergente peut alors être celle exprimée par une telle interprétation (ce serait, par exemple, le désordre qui émerge). Peut-être que cette procédure se limite à la métaphore, il n'en reste pas moins que cette possibilité ajoute aux difficultés de préciser clairement l'explanandum d'une émergence expliquée.

Donc, une émergence expliquée exige l'identification de trois éléments, soit un concept d'émergence qui n'élimine pas la possibilité d'une explication en général, ensuite une stratégie explicative, qui doit être autre qu'une explication de type réductif si l'émergence est définie en opposition à ce type d'explication, et enfin l'explanandum, soit le phénomène d'émergence, dont la description peut être délicate car ce que j'ai appelé les « discontinuités conceptuelles » sont nombreuses, allant de nouveaux concepts incluant des ingrédients inédits dans l'appareil réductif, aux lois-ponts et à l'interprétation de

ces mêmes concepts. Par contre, en coupant la définition de l'émergence de la notion d'explication (généralement de type réductif), comme ce peut être le cas avec l'émergence ontologique²⁹, il est possible d'offrir une explication du phénomène d'émergence.

2.4.3 Émergence qui explique

Le concept d'émergence, on l'a vu, intervient généralement lors d'un échec d'une réduction (épistémologique ou ontologique) et il est donc légitime de poser la question si cet échec implique l'impossibilité (de principe ou en pratique) d'une explication (d'un niveau d'organisation). On peut aussi, à l'inverse, demander si l'émergence elle-même peut expliquer, si elle peut jouer un rôle important dans un explanans.

Une stratégie visant à faire jouer le rôle d'explanans au concept d'émergence dans le but d'expliquer un phénomène donné est de procéder à une « nomicisation » de l'émergence, c'est-à-dire à mobiliser ce concept au sein d'un énoncé nomique, lui-même intégré au sein d'une explication de type déductivo-nomologique. Il a déjà été noté qu'une explication de ce type est compatible autant avec une conception épistémologique des lois, où les lois reçoivent souvent une interprétation instrumentaliste, qu'avec une conception ontologique, comme « loi de la nature ». Cela n'enlève rien toutefois aux problèmes internes que rencontre ce genre d'explication avec la définition d'une loi. Il n'est pas facile d'identifier les desiderata de ce qui constitue une loi et encore moins d'y rattacher un consensus (et ce n'est pas notre objectif ici). À tout le moins, il faut qu'un énoncé soit suffisamment général pour ne pas référer à des objets trop spécifiques, trop spatio-temporellement localisés (par ex. les objets de mon bureau)³⁰. Par voie de conséquence, un énoncé suffisamment général sera en mesure de supporter les conditionnels contrefactuels, soit les énoncés du genre « si les choses avaient été différentes, alors tel état de fait serait vrai » ou « s'il était le cas que *A*, il serait le cas que *B* », où la clause « si les choses avaient été différentes » réfère, entre autres, à d'autres coordonnées spatio-temporelles.

La façon la plus simple, quoique sa justification ne l'est pas, pour donner corps à cette stratégie est sans doute de « réifier », d'« ontologiser » l'énoncé nomique de l'émergence. L'idée de base est donc d'adopter une thèse telle que « il y a une telle chose dans le catalogue ontologique du monde qu'une émergence se produisant dans certaines conditions ». Dans le cas d'une émergence de type

²⁹ On se rappellera toutefois que la stratégie R *infère* l'émergence ontologique de l'émergence épistémologique.

³⁰ Il y a deux caractéristiques supplémentaires généralement attribuées aux lois, complémentaires et souvent antagonistes, soit la « force » et la « simplicité » (Mitchell 1997 ; Carroll 2010).

méréologique, il faudrait alors caractériser les relata et la relation entre ces relata de manière à spécifier une relation ontologique modale, typique de ce que l'on entend par une loi de la nature ou une relation causale. Sans entrer dans le débat animant ce type de caractérisation et de justification des lois, cette relation ontologique modale aurait sans doute la forme d'une relation entre universaux telle que « tout F est G », où F et G seraient deux niveaux de réalité. Le caractère modal de cette relation peut être interprété comme le fait qu'une loi « gouverne » ou « régit » dans une certaine mesure le cours des choses, qu'elle encadre ce qui est naturellement possible. Il peut l'être aussi comme une relation causale, où la base F causerait l'entité émergente G , en sorte que F serait alors ce que Kim (1999 : 13) appelle un « réalisateur » (« *realizer* ») de G . C'est ainsi que cette relation ontologique modale possède un pouvoir explicatif. L'approche ontologique de la mise en œuvre de cette stratégie de « nomicisation » de l'émergence exige donc d'avoir en main un concept d'émergence « positif » en ce sens où il y est décrit, pour ainsi dire, une page du catalogue ontologique du monde.

Ce type de stratégie est possible, dans sa mouture ontologique, avec la proposition de Humphreys (1997) de conceptualiser l'émergence comme une « opération de fusion de propriétés », soit une « opération physique réelle ». Selon cette conception émergentiste, une « fusion » consiste en une sorte d'interaction causale singulière où des instances de propriétés à un niveau inférieur se « combinent » pour donner une instance de propriété d'un tout à un niveau supérieur. En conséquence, les instances de propriétés au niveau inférieur n'existent plus. Une telle opération suffisamment générale peut faire l'objet d'une nomicisation. L'idée générale est donc de dire « il y a une telle chose dans le monde qu'une opération de fusion entre certaines propriétés à certains niveaux » et donc, à la question « pourquoi telle propriété » la réponse est « parce que telle opération de fusion, donc telle émergence ». C'est la raison pour laquelle O'Connor et Wong (2005) exigent que la relation entre la base émergente et ce qui émerge soit causale (ce qu'on peut appeler l'*émergence causale* ; voir aussi Wong 2006). Cela dit, qu'il y ait des relations causales n'est pas difficile à concevoir, bien que certains aient récusé l'existence de la causalité et que sa définition fasse encore l'objet d'intenses débats. Mais qu'il y ait réellement des « opérations de fusion » semble encore moins susceptible d'entraîner l'adhésion. En fait, il s'agit d'une description de la même situation, décrite par l'approche interventionniste, où différents modes d'accès donnent différentes instances de propriétés (celles du niveau inférieur et celle du niveau supérieur), mais où l'on imagine une opération à portée ontologique reliant ces instances de propriétés. Demeure donc

cette perplexité quant à son effectivité, à sa réelle instanciation dans le monde au-delà de sa prétention ontologique³¹.

Le problème avec cette stratégie est donc qu'il n'y pas, ou très peu, de proposition d'un concept « positif » d'émergence, mais plutôt des conceptions de nature épistémologique, comme l'impossibilité, en pratique ou en principe, d'une explication réductive. Plusieurs concepts d'émergence décrivent ainsi une absence, de réduction ou d'explication. Une telle conception ne correspond pas à la structuration effective des processus naturels, ni ne représente une entité avec un quelconque pouvoir causal ou une finalité propre. C'est en ce sens qu'il s'agit d'un concept « négatif ». Mais on conçoit bien que dans une situation où il est admis que la base d'émergence ne puisse expliquer le phénomène émergent le simple fait d'accoler à cette situation le nom « émergence » ne possède en soi aucun pouvoir explicatif. Baptiser notre ignorance ne change en rien ses origines. Que ce concept permette de penser que tel phénomène se réalise, qu'il puisse décrire une situation, celle d'une impossibilité épistémique, ne lui confère aucun pouvoir explicatif.

Hormis cette possibilité de faire appel à ce qu'on a appelé une « nomicisation ontologique », existe une approche s'appuyant sur les théories scientifiques mobilisées dans la description d'un phénomène prétendument émergent. Malgré un caractère émergent défini selon l'échec d'une explication réductive, ce phénomène peut être dit expliqué au sens où il est situé, d'un point de vue théorique, dans un contexte plus large. Il s'agit ici, tel que discuté, de considérer le concept d'émergence d'un point de vue méthodologique : malgré des théories scientifiques et des modes d'accès fiables, on se trouve dans l'impasse (en pratique ou en principe) d'en offrir une explication satisfaisante (généralement, mais pas nécessairement, respectant le canon réductionniste). Par exemple, le phénomène de la vie est largement inexplicable, au sens réductionniste du terme, par la physique, même si l'on ne doute pas de l'existence d'organismes vivants. L'explication particulière que peut apporter le concept d'émergence vient de la mise en lumière d'une situation théorique où est affirmée l'existence de certaines entités à des niveaux d'organisation différents, malgré l'impossibilité d'expliquer les unes par les autres et malgré la confiance que l'on peut avoir en leur existence et leurs relations mutuelles. Il ne s'agit donc pas d'une explication du phénomène qualifié d'émergent, mais plutôt de l'explication des défis théoriques qu'il pose (ou impose).

³¹ En revanche, si cette approche est avérée, s'il y a vraiment « fusion », alors il serait possible d'offrir une explication de l'émergence en général (et pas seulement des instances d'émergence), en autant que l'on dise pourquoi il existe de telles opérations de fusion.

2.5 Conclusion

Il est naturel devant la complexité des phénomènes, mais aussi des théories qui en rendent compte, de procéder à son analyse, de réduire le tout aux propriétés des parties. Car une analyse se définit comme une opération par laquelle l'esprit décompose un ensemble constitué, pour en déceler l'autonomie des parties, pour en apprécier mieux la congruence ou la finalité, ou simplement pour rendre accessible chacun de ses éléments. L'aboutissement de cette analyse est donc bien souvent une explication du phénomène complexe, à tout le moins on est en droit de s'attendre à une meilleure compréhension. En supposant que cette explication reprenne les concepts associés aux parties, bref au résultat de l'analyse, il s'agira d'une réduction ou d'une explication réductive. Mais il arrive bien entendu qu'une simple analyse ne soit pas suffisante pour comprendre un phénomène et qu'une explication réductive soit impossible. Cependant, il faut préciser les termes de l'analyse et donc de l'explication, les entités et les théories en jeu.

Car il y a peu de notion aussi centrale mais aussi polysémique que celle d'explication. Centrale, car elle peut jouer un rôle normatif en participant à un critère de démarcation entre science et non-science (donc un travail pour le philosophe). Polysémique, car plusieurs modèles ont été proposés par les philosophes et un regard attentif à la pratique scientifique révèle aussi un pluralisme. Tout n'explique pas pour autant et c'est pourquoi j'ai soutenu un pluralisme restreint. Car les critiques des théories orthodoxes de l'explication ont surtout porté non pas sur la pertinence de l'approche, mais plutôt sur leur nécessité étant donné les exigences qu'elles placent sur l'explication. C'est-à-dire qu'un explanans exprimés selon les termes de ces modèles, comme une loi ou une cause, serait suffisant, dans des circonstances particulières, mais pas nécessaire en toutes circonstances à une « bonne » explication, satisfaisante. D'autres stratégies explicatives ont été proposées, mais un critère nécessaire, sans être suffisant, semble être celui de l'expectabilité, comme quoi l'on doit être en droit de s'attendre que le phénomène devant être expliqué se produise (dans des circonstances données). Une manière de préciser ce critère est de faire appel aux « schèmes systématiques de dépendance contrefactuelle » (Woodward 2003a), qui peuvent fournir des réponses à des questions contrefactuelles, de dire en quoi l'explanandum serait différent si les choses avaient été différentes. Mais analyser la distinction profonde entre l'expectabilité et ces schèmes de dépendance contrefactuelle demanderait de plus amples développements.

Afin d'interpréter l'expectabilité comme critère d'une bonne explication, j'ai ensuite proposé la modalité (au sens faible) et la naturalité. Selon la modalité une explication doit montrer que ce qui s'est passé, et qui doit être expliqué, « devait » se produire. Ce critère peut se représenter par le concept de

d'espace des phases, qui permet de définir deux classes d'énoncés contrastives définissant ce qui est possible en opposition à ce qui ne l'est pas. Dans des circonstances données, donc, l'explanandum est d'abord situé parmi cette gamme de possibles (qui offre déjà une explication sommaire) et ensuite il est davantage circonscrit par un lien modal définissant une trajectoire dans l'espace des phases, comme une cause ou une loi stricte, ou selon un lien plus faible avec une loi probabiliste. La naturalité, par la subsumption d'un concept sous une classe de référents similaires ou familiers, permet d'éliminer, à l'instar d'une gamme de possibles, les explanans incompatibles avec un certain arrière-plan ontologique, faisant intervenir, par exemple, des divinités ou des êtres invisibles. Bien entendu, le domaine de validité d'une loi circonscrit aussi, partiellement du moins et sans nécessairement faire référence à ce genre d'arrière-plan, les entités pertinentes à l'explication. Un autre exemple pouvant être interprété par le critère de naturalité est celui d'une explication où il est montré que l'explanandum est un cas particulier, un exemple, d'une loi, sans pourtant le déduire (Forge 1980). Une « bonne » explication peut donc se présenter en divers degrés de force et doit montrer que ce qui s'est passé (explanandum) pouvait (gamme de possibles) se passer, plutôt que tel autre explanandum (impossible) ; ou que cela devait, selon une modalité faible, se passer ; ou que dans des circonstances différentes cela se serait aussi, naturellement, produit.

Ce cadre définitionnel de l'explication permet de faire face aux défis que posent les systèmes complexes avec de nombreux niveaux d'organisation. Un niveau d'organisation est une description d'un ensemble de propriétés et de régularités suffisamment stable pour être déterminées de manière robuste selon les termes de Wimsatt. Cette conception des niveaux d'organisation est compatible, sans l'impliquer, avec une approche causaliste des propriétés et n'adopte pas nécessairement un fondationnalisme ontologique où des entités élémentaires à un niveau fondamental posséderaient des propriétés intrinsèques, tandis que des entités de niveau supérieur présenteraient plutôt des propriétés structurales. Conséquemment, un niveau d'organisation peut être considéré comme un « niveau de réalité », entendu qu'il s'agit des propriétés des entités que le constituant et qui se révèlent par des modes d'accès généralement empiriques, à partir desquels est décrit l'ensemble de ces propriétés interprétées par nos théories scientifiques, alors que ces théories sont aussi, en partie, constituées par le biais de ces modes d'accès. En effet, le cadre réaliste adopté soutient la thèse que le monde est connaissable sans pour autant l'harnacher à l'exigence de la vérité absolue des théories.

Il n'est donc pas question d'endosser un réductionnisme global où seul le domaine physique aurait la légitimité du réel, ni un pluralisme défendant la différence intrinsèque ou l'absence d'influence entre divers niveaux d'organisation. Pour autant, un réductionnisme plus modeste, local, est souhaitable, et compatible avec la conceptualisation des niveaux d'organisation précédente. L'émergence peut donc aussi être considérée comme un principe méthodologique forçant la recherche d'explications

réductionnistes autant que possible en reconnaissant leurs forces et leurs faiblesses. Affirmer qu'un phénomène est émergent c'est affirmer qu'il existe un lien modal (par ex. la survenance) entre des entités à des niveaux d'organisation différents sans qu'une explication réductive ne soit possible (par ex. il n'y pas de température sans atome). En ce sens, identifier un cas d'émergence revient à caractériser (et possiblement l'expliquer) une situation théorique particulière où différents modes d'accès permettent de définir des niveaux d'organisation sans qu'il n'y ait de réduction totale (identité ou élimination). L'émergence, son phénomène, peut donc être expliquée puisqu'une explication réductionniste demeure possible si l'émergence n'est pas synonyme de non-explicabilité. Il faut alors faire preuve de prudence afin de ne pas opérer un glissement de sens d'un cas d'émergence expliquée à celui d'une émergence qui explique. Dans ce cas, la stratégie explicative doit être clairement définie et il n'est pas nécessaire qu'elle soit de type réductif. Concevoir l'émergence comme le terme d'un processus ou le processus lui-même (par ex. une « fusion ») permet aussi de la mobiliser au sein d'une explication du l'explanandum (émergent).

3 Historique du concept d'entropie

RÉSUMÉ. Le concept d'entropie est apparu dans le contexte de la révolution industrielle du XIX^e siècle. Il a d'abord été pressenti dans les travaux de Sadi Carnot, lesquels ont donné le concept de cycle réversible idéalisé où une quantité de chaleur est maximale convertie en travail. Il a ensuite été baptisé et formalisé par Rudolf Clausius, le plaçant alors au centre du second principe de la thermodynamique exprimant l'impossibilité du mouvement perpétuel. Un changement de paradigme, opéré par Ludwig Boltzmann et Josiah Willard Gibbs, d'obtenir une formulation de l'entropie en fonction de quantités statistiques sur de grands ensembles de molécules régies par la mécanique classique. Sans doute en raison de son caractère « prodigieusement abstrait » (Poincaré), il souffre de polysémie et il a été récupéré dans diverses disciplines, de la génétique des populations à la théorie de l'information. Cependant, malgré les nombreuses interprétations et l'ambiguïté dont il est l'objet, le concept d'entropie demeure un concept fondamental, tandis qu'on lui prête des ambitions grandioses comme celle d'expliquer la directionnalité du temps.

Eighty percent of success is showing up.

Woody Allen

Le second principe de la thermodynamique est souvent présenté comme l'une des plus importantes lois de la science moderne ou comme l'une des lois que tout individu « scientifiquement informé » devrait savoir. Il en va de même du principe entropique, censé le recouvrir avec un concept représentant une quantité physique, l'entropie. Ce concept est apparu dans le contexte de la révolution industrielle du XIX^e siècle et s'inscrivait alors dans un effort visant à améliorer l'efficacité des machines thermiques. Plusieurs travaux importants de la physique de l'époque se penchaient en effet sur les inévitables limitations de rendement des machines thermiques, puis, plus globalement, sur diverses transformations associées à la chaleur et au travail mécanique, des problèmes dont la résolution allait aboutir au concept d'énergie. C'est pourquoi ce concept fondamental a toujours été dans une position hybride entre la physique et l'ingénierie. Car il exprime non seulement une idée fondamentale en physique, et on pourrait dire plusieurs idées fondamentales dont la clarification exige un traitement étendu (Chapitres 4 et 5), mais aussi une incapacité technologique, à savoir l'impossibilité du mouvement perpétuel de second type (discuté plus loin). C'est aussi pourquoi il est intimement lié au concept d'énergie avec lequel il forme les fondements de la thermodynamique classique dont l'élaboration a été pour le moins tortueuse.

Conceptuellement, le concept d'entropie est défini à partir des concepts de chaleur et de température. Il est dérivé formellement au moyen d'idéalisations définissant un cycle réversible au sein duquel certaines valeurs physiques se compensent mutuellement, ce qui permet de définir une fonction d'état. Bien que son rôle soit très important, en ceci qu'il est mis à contribution dans plusieurs explications et qu'il participe à la définition de certains concepts aussi importants que la température, sa dérivation, sa preuve, demeure objet de débats, quand elle n'est pas simplement occultée, et c'est pourquoi on ne présente souvent que sa définition à la manière d'un postulat. Sans statuer ici sur la validité de cette stratégie générale reposant sur un cas limite idéalisé jamais instancié (Chapitre 4), on peut se poser la question si elle ne découle pas, en peu de mots, de la conciliation du desideratum d'une fonction d'état et d'un principe stipulant une impossibilité, celle du mouvement perpétuel. Quoiqu'il en soit, ce cas limite idéalisé est mis en opposition contrastive au comportement de systèmes susceptible d'être corroboré et dont l'entropie augmente. Il y aurait ainsi une tendance à l'augmentation d'entropie exprimée par le principe entropique. L'augmentation apparemment inévitable d'une quantité physique exemplifie une irréversibilité. Or, l'irréversibilité exprimée par le principe entropique est en contradiction avec la réversibilité des lois de la mécanique classique. Cette contradiction est particulièrement problématique si l'on pose l'hypothèse, à première vue anodine, que les systèmes auxquels on attribue une valeur d'entropie sont constitués de plusieurs particules soumises à ces lois réversibles.

Ce problème théorique est persistant dans l'entreprise visant à retrouver, à prédire, à expliquer le comportement des systèmes de la thermodynamique à partir des variables et des lois de la mécanique classique. Cette entreprise est née à la fin du XIX^e siècle alors que l'hypothèse atomiste, quoique fortement contestée, faisait son chemin, tout comme la méthode appliquant les statistiques et les probabilités aux phénomènes naturels. D'où la mécanique statistique, qui applique les méthodes statistiques et la mécanique classique aux très nombreux constituants idéalisés de la matière, soit des modèles d'atomes et de molécules comme des sphères parfaites, dans le but de rendre compte des propriétés thermodynamiques des grands systèmes composés de ces constituants. L'inaccessibilité empirique de ces derniers, par leur taille individuelle mais aussi par le nombre gigantesque, est en grande partie responsable du besoin d'en appeler à des hypothèses qui peuvent bien paraître rationnelles mais qui, justement, ne peuvent bénéficier de la caution de l'expérience. La version statistique du concept d'entropie est ainsi traversée par ses hypothèses, avec ce que cela comporte de problèmes épistémologiques. Il n'est donc pas surprenant qu'il fasse l'objet d'une polysémie prenant la couleur de ces hypothèses.

L'impact du concept d'entropie au sein cette entreprise théorique d'une part, avec ses échecs et ses succès, mais aussi en thermodynamique d'autre part, est ainsi considérable pour la philosophie des

sciences. Car non seulement il propose une assise à l'explication des phénomènes irréversibles et même, comme plusieurs l'ont suggéré, à la directionnalité du temps, mais il offre aussi l'occasion d'une discussion fort pertinente sur les relations inter-théoriques, sur les niveaux d'organisation de la matière et, sans l'oublier, sur l'application du concept de probabilité dans les sciences naturelles. Un examen critique est donc de mise et un exposé historique de ses fondements et développements est loin d'être superflu. Le besoin d'une analyse critique exige en même temps une position stable à cette même analyse. Cette position est sans doute introuvable, à moins de se satisfaire de l'évidence cartésienne, et ne peut certainement pas être elle-même à l'abri de l'examen critique. Mais cela étant dit, l'intérêt d'une analyse historique apparaît ici comme la possibilité de retrouver ce qui est à la base de la pérennité du concept d'entropie *malgré* cette polysémie. À l'instar de l'idée de la sélection naturelle ayant reçu plusieurs interprétations tout en restant théoriquement indispensable (ou presque), il y a sans doute une idée tout aussi forte avec le concept d'entropie que l'on associe à l'impossibilité du mouvement perpétuel. Il est effectivement possible de rendre compte de certaines de nos intuitions les plus fortes au moyen de ce concept. Mais, s'il va de soi que l'on ne peut se limiter à ces intuitions dans l'examen conceptuel critique qui nous intéresse ici, il n'en demeure pas moins qu'elles peuvent révéler quelque chose de fondamental à l'œuvre dans l'histoire des développements conceptuels de l'entropie. Il faut alors remonter à l'époque où les chaudières au charbon commençaient à s'implanter en Europe, semant l'illusion que l'homme allait bientôt s'affranchir du travail manuel.

3.1.1 Révolution scientifique

Si le concept d'entropie est apparu dans le contexte de la révolution industrielle du XIX^e siècle, ce contexte doit d'abord être mis en lumière par celui d'une autre révolution, la révolution scientifique. Il ne fait aucun doute que la notion de révolution en histoire fasse débat. Elle évoque une radicale discontinuité mais aussi une sorte de directionnalité, d'irréversibilité, un impossible retour en arrière. C'est pourquoi l'objet même de la discipline historique, le passé, semble en rupture méthodologique avec les sciences dites naturelles exigeant la récursivité. Pourtant, l'explication historique ne peut faire l'économie d'une certaine continuité, à préciser toutefois, comme dans une explication exemplifiant une régularité causale. Cette dichotomie continuité-discontinuité est donc centrale en histoire, sans aucun doute irréductible à l'un de ses termes, et l'insistance sur l'un d'eux conditionne fortement la position que l'on peut avoir sur une occurrence possible de ce que l'on considère comme une « révolution ». Le cas de l'histoire des sciences n'est pas moins compliqué, au contraire, et ce concept de « révolution scientifique » est l'objet de débats largement conditionnés par cette dichotomie. Et ces derniers sont d'autant plus pressants qu'ils alimentent, tout en en subissant l'influence, les débats de la

philosophie des sciences. Car la première ne saurait se passer d'outils critiques et la seconde serait vide sans exemples tirés de l'histoire, ancienne et récente. On peut déjà remarquer qu'il n'est plus d'usage courant de parler de l'histoire de *la* science, qui serait une, et dont le devenir serait continu et uniforme. Celui-ci apparaît plutôt irrégulier, affichant une diversité de concepts, de méthodes et de résultats, difficilement conciliable avec une théorie de la connaissance unificatrice. L'histoire de *la* science paraît donc un projet « idéologique » (Fichant & Pécheux 1969 : 96). Si ce projet n'est pas, *a priori*, illégitime et dénué d'intérêt, il exige en revanche un discours justificateur allant bien au-delà de la présente discussion. Ces nuances faites, il m'est permis de présenter, brièvement certes, ce que j'entends par révolution scientifique¹.

Sans en apporter une preuve solide, on peut d'abord noter que l'époque du tournant du XVII^e siècle était en effet soumise à un processus de changement, comme en témoigne l'utilisation fréquente du mot « nouveau » dans le titre de certains ouvrages, dont l'*Astronomia Nova* (1609) de Képler, le *Novum organum* (1620) de Bacon et les *Deux nouvelles sciences* (1632) de Galilée. Cette époque est celle qu'on associe à la révolution scientifique, au gré de circonstances particulières, comme l'avènement de l'imprimerie européenne, la redécouverte des Amériques et la sécularisation (Butterfield 1949 ; Watson 2005). Mais l'objectif ici n'est pas d'expliquer ses origines mais plutôt de caractériser ses legs : pourquoi et en quoi cette révolution représenterait-elle la « plus profonde révolution de la pensée humaine depuis la découverte du Cosmos par la pensée grecque » (Koyré 1966 : 166) ? La question est évidemment difficile puisqu'elle revient à caractériser, ici aussi, la « science moderne » et, par le fait même, un genre d'idéal normatif auquel devraient se soumettre les entreprises d'explication du monde. Ce faisant il convient d'explicitier les deux termes de la transition, voire de la rupture, entre ce qu'il y avait avant et ce qu'il y avait après (au même titre que la révolution française ne fait guère sens sans un exposé de ce qu'est la monarchie). Ici, la science aristotélicienne, première influence du savoir scholastique, fait figure de ce qui est révolu. Mais déjà plusieurs difficultés se présentent car les traits distinctifs recherchés peuvent difficilement être trouvés dans les *objets* de ces savoirs, après tout le savoir antique traitait aussi de la chute des corps, ni dans ses *produits*, si tant que est que la cohérence logique de nos théories n'est pas une nouveauté et que les concepts scientifiques contemporains sont historiquement situés, donc appelés à changer. En ce sens, il convient de caractériser plutôt une *méthode*, même si, bien évidemment, plusieurs nouveaux concepts, théories et lois, plusieurs nouveaux « paradigmes » selon l'expression consacrée de Kuhn (1962), ont été proposés. Cette méthode, loin d'avoir été

¹ Même si cette question est vaste et mérite de plus amples développements, je crois utile et même nécessaire de présenter ce qui caractérise cette période qui, quoiqu'on puisse en penser, a façonné ce qu'on appelle la « science moderne ». La rédaction de ce chapitre a été grandement influencée par le cours d'histoire des sciences et des techniques modernes que j'ai donné de 2008 à 2014 à l'École Polytechnique de Montréal.

universellement appliquée, s'est graduellement imposée suite aux présentations partielles mais complémentaires du *Novum Organum* (1620) et de la *Logique de Port-Royal* (1662).

La présentation des grandes lignes de cette méthode dite scientifique gagne en clarté si on l'encadre par une distinction entre « contexte extra-théorique », le processus *de facto*, qui concerne les facteurs sociaux et économiques d'un fait ou d'un événement de l'histoire des sciences comme la production d'une théorie, et « contexte intra-théorique », la défense *de jure* d'une théorie, qui concerne ses preuves². Dans le premier cas, on peut noter les phénomènes d'institutionnalisation et de professionnalisation, c'est-à-dire la transformation du philosophe au scientifique, du scribe au savant, ce passage du « savant » isolé et non rémunéré, souvent soumis au mécénat, dilettante et autodidacte, dont les recherches se faisaient souvent « dans le secret », au chercheur professionnel rémunéré, avec une formation théorique institutionnalisée, devant « valoriser » les produits de ses recherches, par des publications ou l'obtention de propriétés intellectuelles³. Dans le second cas, on fait face à tout un programme pour la philosophie des sciences. En bref, il s'agit dans un premier temps de l'expérimentation et de l'instrumentalisation, c'est-à-dire, en substance, à cette transformation de l'investigation des phénomènes naturels, ce passage d'une activité spontanée d'observation de certains faits, tels qu'ils se présentent à nous d'eux-mêmes, sans que nous les ayons provoqués, à l'observation provoquée volontairement, d'une façon systématique et dans des conditions contrôlées, avec le maximum de précision, à l'aide d'instruments de mesure généralement issus de connaissances théoriques précises⁴. Dans un second temps, la formalisation mathématique des phénomènes naturels est un facteur primordial dans l'évolution du savoir humain, une méthode sans doute initiée par

² Cette distinction est semblable mais tout de même différente de celle, bien connue, entre *contexte de justification* et *contexte de découverte* (Ladyman 2007; Schickore 2014).

³ On note aussi l'évolution dans les institutions d'enseignement : de l'université axée sur la formation et des académies axées sur l'avancement des sciences au modèle allemand initié par Wilhem von Humbolt avec l'université de Berlin en 1809 où l'on tente de concilier recherche et enseignement ; diffusion de ce modèle dans l'Europe occidentale et en Amérique du Nord au XIX^e siècle ; apparition des instituts de recherche publics et privés, des chercheurs fonctionnaires et de la troisième « vocation » de l'université avec la valorisation des produits de recherche, au milieu du XX^e siècle ; apparition de l'économie du savoir et des chercheurs entrepreneurs à la fin du XX^e siècle. Les États prennent progressivement en charge les sciences : par ex., la *Royal Society of London* obtient une charte officielle du roi Charles II (1662) ; financement étatique par le ministre Colbert de l'*Académie des sciences de Paris*. Le scientifique commença à recevoir un salaire mais aussi des responsabilités : par ex., des pressions sont faites pour que la *Royal Society* examinent les demandes de brevet et l'État impose aux membres de l'*Académie* certaines tâches telles que le tracé de cartes géographiques. Apparition d'écoles d'application (surtout en France) : par ex., École des ponts et chaussées (1744), École militaire de Paris (1756), École polytechnique (1794). Les périodiques scientifiques, notamment le *Journal des Sçavans* de l'Académie des sciences et les *Philosophical Transactions* de la Royal Society, prennent de l'ampleur à cette époque.

⁴ L'utilisation d'instruments dans la pratique scientifique s'est surtout intensifiée au début du XVI^e siècle. Entre autres, on voit chez Galilée l'utilisation du télescope, bien sûr, mais aussi de plans inclinés, de thermoscopes et de chronomètres, plutôt rudimentaires dans ce dernier cas mais tout de même novateur dans leurs applications. Sans les mettre véritablement en pratique, Bacon définit les conditions d'accomplissement de la tâche collective de construction de connaissances par une démarche critique, partant du simple constat sensoriel et s'élevant par induction à des propositions de portée générale. Les fondateurs de la *Royal Society of London* ont d'ailleurs revendiqué l'héritage de cette démarche inductive et expérimentale.

Archimède mais remise de l'avant par Galilée, avec son « livre de la nature » qui serait écrit en langage mathématique, et enfin systématisée par Newton. Ce dernier, ayant proposé une méthode qui allait influencer, peu ou prou, toute la science du XVIII^e siècle, présentait aussi ses « principes mathématiques » tout en insistant sur la « philosophie expérimentale ». Ce subtil amalgame de mathématisation et d'expérimentation est sans doute ce qui caractérise le plus, de l'intérieur, la méthode scientifique moderne⁵. (Pour des discussions sur la révolution scientifique, voir Butterfield 1949 ; Blay & Halleux 1998 ; Gingras, Keating & Limoges 1998 ; de la Cotardière 2004 ; Mazaauric 2011.)

Sans en présenter toutes les nuances, et à l'instar d'un contexte théorique guidant l'application de concepts, cette brève présentation dresse le portrait plus large de l'évolution de la pensée scientifique dans lequel s'inscrit, plus spécifiquement, l'élaboration ainsi que l'évolution du concept d'entropie. Car ce concept fondamental est au carrefour d'un ensemble de débats épistémologiques capitaux impliquant différentes visions de ce que devraient être la science et son arrière-plan métaphysique.

3.1.2 Mise en contexte historique du concept d'entropie

Au tournant du XIX^e siècle, l'Angleterre avait acquis une position plus qu'enviable sur la scène du commerce international grâce au développement de machines à vapeur qui avaient ranimé ses mines de houille tout en justifiant leur exploitation par leur consommation. L'extraction effrénée de ce combustible était commandée par la demande croissante en puissance des industries sidérurgique, textile et minière. Il était devenu commun de voir les imposants balanciers de ces machines opiner de jour comme de nuit tandis que le paysage urbain se transformait : l'Angleterre créait une révolution aussi bien qu'elle la subissait. Cela n'était pas étranger au fait qu'elle avait harnaché la chaleur à son profit, c'est-à-dire en exploitant la densité d'énergie issue de la combustion des ressources fossiles afin de produire de la vapeur comme force motrice. Et l'ingénieur français Sadi Carnot (1796-1832) le savait : sa principale motivation résidait ainsi dans la compréhension des lois régissant le fonctionnement des « machines à feu », afin de les améliorer et ainsi pouvoir concurrencer les Anglais dans leur hégémonie industrielle⁶. Il commença d'ailleurs son ouvrage majeur par évoquer le potentiel

⁵ Par exemple : « il faut retenir que l'originalité de Galilée a [...] été de combiner les méthodes géométriques d'Archimède et l'expérimentation systématique » (Gingras, Keating & Limoges 1998 : 248).

⁶ L'influence de l'Angleterre (ou de la Grande-Bretagne en général), comme modèle économique et industriel, doit toutefois être nuancée, car l'industrialisation européenne doit être vue moins comme la répétition d'un modèle que comme un processus complexe. Bien que la France accusait un retard industriel sur la Grande-Bretagne, en raison principalement de son faible taux de natalité et de l'accessibilité réduite au minerai de charbon, l'obligeant ainsi à se tourner vers la puissance hydraulique, sa croissance économique au tournant du XIX^e siècle fut similaire à celle de sa concurrente (voir Cameron 1985).

énorme de l'énergie calorifique, soit la chaleur. L'étude de ces machines était donc « du plus haut intérêt » car elles lui paraissaient « destinées à produire une grande révolution dans le monde civilisé » (Carnot 1824 : 2). C'est pourquoi les progrès réalisés au XIX^e siècle dans l'étude de la chaleur et du travail sont indissociables des progrès techniques réalisés depuis le début du XVIII^e siècle dans la mise au point de machines de puissance (Kuhn 1959 ; Cardwell 1967 ; Prigogine & Stengers 1979)

À cette époque, la nature de la chaleur était plutôt mal comprise et deux interprétations se faisaient concurrence : l'interprétation *substantielle*, qui postulait l'existence d'un « fluide impondérable » étant la cause de la chaleur voire la chaleur elle-même, et l'interprétation *mécaniste*, selon laquelle la chaleur n'était qu'un mode de mouvement des particules de la matière ordinaire. L'interprétation dominante était alors celle de la théorie substantielle, qui présentait toutefois plusieurs versions et qui était principalement soutenue par les chimistes, dont Antoine Laurent de Lavoisier (1743-1794) qui a donné le nom de « calorique » à ce prétendu fluide et le plaçant dans la classe des « corps simples », sans doute en raison de leur plus grand intérêt à la « diversité qualitative de la matière » (Vassails 1950 : 224). Plusieurs observations semblaient corroborer cette interprétation : le mélange de deux substances, l'une chaude et l'autre froide, menait semble-t-il à l'équilibre des températures en raison du transfert de calorique d'une substance à l'autre ; le martelage d'une barre de fer expulsait le calorique et augmentait ainsi la température de cette même barre ; la combinaison du calorique avec les particules de matière augmentait la fluidité de celle-ci en même temps que sa température (Holton & Brush 2001 ; Chang 2004). Néanmoins, parallèlement, l'interprétation mécaniste avait fait quelques adeptes, dont Francis Bacon (1561-1626), René Descartes (1596-1650) et Robert Boyle (1627-1691), lesquels insistaient plutôt sur la forme et le mouvement des particules matérielles pour rendre compte des phénomènes de composition ou de décomposition des corps. Cependant, au tournant du XIX^e siècle, la situation s'apparentait à un « agnosticisme généralisé » (Chang 2004 : 74) résultant en un relatif désintérêt dans l'investigation théorique sur la nature de la chaleur.

Cette valse hésitation peut s'expliquer, d'une part, par des manques conceptuels (comme une distinction claire entre chaleur et température) et, d'autre part, par l'absence de méthodes fiables de thermométrie et de calorimétrie (Holton & Brush 2001). Or, la théorie substantielle reposait sur deux hypothèses dont l'une allait lui être fatale, celle voulant que le calorique ait un poids et celle de sa conservation. L'objection la plus célèbre est venue de Benjamin Thomson, comte de Rumford (1753-1814), qui dirigeait un atelier de forage de canons à Munich. Il avait observé que le travail rapide des métaux par machines-outils conduisait à la production de quantités énormes de chaleur, particulièrement lorsque les alésoirs étaient émoussés ; or, le frottement pouvait être poursuivi indéfiniment et ainsi produire de la chaleur en quantité illimitée ; ce que ne parvenait pas à expliquer la théorie substantielle selon laquelle le calorique se conservait. Une expérience similaire, proposée par

Boyle et effectuée par Humphry Davy (1778-1829), montrait que deux glaçons fondaient plus rapidement lorsque frottés ensemble. Une complication supplémentaire est venue de l'association entre lumière et chaleur. En effet, la chaleur rayonnée (rayonnements infrarouges) présentait des propriétés similaires à celles de la lumière, mais tant que cette dernière était considérée comme une émission de particules, il était difficile d'établir un rapprochement. Toutefois, grâce aux succès des travaux de Thomas Young (1773-1829) et Augustin Fresnel (1818-1827) présentant la lumière selon une interprétation ondulatoire, il était alors plus facile de dégager une similitude complète de comportement, comme la réflexion et la polarisation, ce qu'a fait explicitement André-Marie Ampère (1775-1836), qui n'a malheureusement pas été suivi en France.

Des éléments étaient donc présents pour apporter une nouvelle conception de la chaleur en l'associant à ce qui est maintenant appelé l'énergie, bien qu'à cette époque ce terme était vaguement associé à ceux de « force », de « cause » et de « puissance ». Or, ce contexte de révolution industrielle, aux prises avec d'insistants problèmes d'ordre technologique, tout ce qu'il y a de plus concrets, allait justement permettre d'assener le coup de grâce à la théorie du calorique en établissant l'équivalence entre la chaleur et le travail. Comme l'exprime Vassails (1950 : 231) :

[...] dès les dernières années du XVIII^e siècle Watt et Boulton pratiquait l'équivalence sans le savoir puisqu'ils calculaient combien de livres d'eau l'on pouvait élever d'un pied de haut, de boisseaux de blé l'on pouvait moudre, de broches de coton l'on pouvait faire tourner pendant une heure, etc., en brûlant un boisseau de charbon. Ils avaient, d'une manière toute empirique, dégrossi l'économie de la machine, poussés par la nécessité de la vendre, de convaincre pour cela les acheteurs notamment les exploitant miniers. Mais désormais il n'était plus possible de perfectionner donc de comprendre clairement cette économie sans prendre conscience que la chaleur et la « force » mécanique peuvent se convertir mutuellement l'une dans l'autre, selon une règle d'équivalence quantitative.

Cette conversion, et plus précisément celle de la chaleur en travail, est sans doute ce qui caractérise le plus la révolution industrielle, ce qui a permis d'« ouvrir la boîte de Pandore » (Wrigley 2010 : 1), d'augmenter substantiellement le pouvoir de production dont parlait Marx. On peut donc dire que le concept moderne d'énergie est issu, pour l'essentiel, d'une conciliation conceptuelle entre, d'une part, une appréciation de certains phénomènes manifestes dans le fonctionnement des machines selon une règle quantitative, mais aussi, nous allons le voir, d'une sorte d'a priori métaphysique rattaché aux principes conservatifs.

3.1.3 Le concept d'énergie

L'idée d'énergie est beaucoup plus vieille que son nom. Chez les Grecs anciens, et plus spécifiquement chez Aristote, le mot *ενεργεια* (*energia*) renvoie au mode d'être en acte et se distingue de *δυναμικη* (*dunamis*), le mode d'être en puissance. L'*energia* est ainsi une activité en cours, que l'on peut associer au « travail » ou à une « action interne ». Le terme actuel d'énergie possède aussi ces deux connotations de puissance et d'action, mais il a acquis « un sens mathématique précis » (Balian 2013 : 1). Dans la langue française, le terme apparaît semble-t-il avec Jean Bernoulli, le 26 janvier 1717, dans une lettre au mécanicien Pierre Varignon, où il définit de manière différentielle l'énergie comme le produit de la force appliquée à un corps par le déplacement infinitésimal subi par ce corps sous l'effet de cette force, produit désigné plus tard comme le travail mécanique de cette force.

L'idée voulant qu'une seule « force » soit à l'œuvre dans une multitude de phénomènes pourtant divers, bref qu'une grandeur se conservait, a commencé à s'imposer déjà à la révolution scientifique. Ainsi Galilée, dans ce qui allait mener à la modernisation du concept d'inertie, reconnaît-il que la vitesse d'un corps demeure constante en l'absence de force extérieure. Aussi, René Descartes admet que la quantité de mouvement, comme quantité scalaire, se conserve lors du choc de deux corps, tout comme Christiaan Huygens (1629-1695), qui reconnaît également que la somme des quantités mv^2 , la « force vive », se conserve lors d'un choc élastique, ce qu'admettait aussi Wilhem Leibniz (1646-1716). Dans ce contexte de triomphe de la physique newtonienne, l'idée d'interconvertibilité suivait un schème réductionniste de type laplacien se rapportant à une force d'ordre supérieur, en ce sens où la diversité des phénomènes énergétiques étaient censés pouvoir se ramener à une sorte de force de type mécanique. D'ailleurs, tous les problèmes de la mécanique classique qui prévalaient alors pouvaient et peuvent encore être résolus sans le concept d'énergie. Mais il manquait la notion énergétique d'un « réseau d'équivalences comptables entre les grandeurs mesurables dans les domaines en interaction » (Merleau-Ponty 1979 : 318). À cet égard, les ingénieurs ont joué un rôle majeur en insistant sur les notions de chaleur, de travail et d'efficacité, relativement éloignées du domaine de la dynamique newtonienne. En effet, la réponse se trouvait dans l'idée que les « forces », un concept alors imparfaitement élucidé, à l'œuvre dans les phénomènes naturels sont multiples et diverses, mais qu'il existe entre elles tout un réseau d'interconnexions et qu'elles peuvent se convertir mutuellement.

Dans son article désormais classique, Kuhn (1959) soutient la thèse selon laquelle le concept moderne de l'énergie constitue un cas de « découverte simultanée » et identifie douze pionniers⁷ qui,

⁷ Julius Robert Mayer, James Prescott Joule, Ludwig August Colding, Hermann von Helmholtz, Sadi Carnot, Marc Séguin, Karl Holtzmann, Gustave-Adolphe Hirn, Carl Friedrich Mohr, William Robert Grove, Michael Faraday et Justus von Liebig.

entre 1842 et 1847, ont énoncé l'hypothèse de la conservation de l'énergie, laquelle ne serait ni plus ni moins que la contrepartie théorique à quatre décennies de découvertes de laboratoire. Il propose trois facteurs prépondérants dans l'émergence d'éléments expérimentaux et conceptuels essentiels à l'élaboration de ce concept : (i) la *disponibilité des processus de conversion*, (ii) l'*intérêt pour les machines* et (iii) la *philosophie de la nature*. La discussion de ces trois facteurs sera suivie d'une brève présentation de trois pionniers du principe de la conservation de l'énergie à l'origine du premier principe de la thermodynamique.

En ce qui concerne le premier facteur, la conversion du mouvement en chaleur (frottement) et celle de la chaleur en travail (machine à vapeur) étaient bien connues, de même que les relations entre réactions chimiques et chaleur ; mais l'invention en 1800 de la pile électrique par Alessandro Volta (1745-1827) a accéléré cette « disponibilité » : on était alors en mesure de prouver qu'une réaction chimique pouvait produire de l'électricité et, peu de temps après, William Nicholson (1753-1815) et Anthony Carlisle (1768-1840) démontrèrent la transformation inverse en réalisant l'électrolyse de l'eau ; la découverte en 1820, par Hans Christian Oersted (1777-1851), de l'effet magnétique du courant électrique ouvrait la voie à André-Marie Ampère (1775-1836) qui établit la connexion entre électricité et mouvement, le magnétisme étant électrodynamique ; peu avant Michael Faraday (1791-1867) qui découvrit l'induction électromagnétique, à l'origine des moteurs électriques et donc de la conversion de l'électricité en mouvement ou en travail. Vers 1830, la physique expérimentale tirait ainsi une conclusion qui allait mener à la conception moderne de l'énergie, à savoir que les forces à l'œuvre dans les phénomènes naturels sont « multiples et diverses, mais qu'il existe entre elles tout un réseau d'interconnexion, et qu'elles peuvent se convertir les unes dans les autres » (Merleau-Ponty 1979 : 316). La notion qualitative de l'interconvertibilité a donc joué un rôle heuristique important dans la formulation du concept mais elle fut appuyée par celle de l'équivalence numérique et comptable : les diverses mesures des effets interconvertibles étaient quantitativement constantes. Il a donc fallu passer d'une notion *qualitative* à une notion *quantitative*, il a fallu passer de la *conversion* à la *conservation*.

2. Le deuxième facteur s'inscrit évidemment dans le contexte de la révolution industrielle du tournant du XIX^e siècle. À cette époque, l'ingénierie européenne tentait de contribuer à l'essor industriel et son enseignement se transformait au gré des bouleversements politiques et du positivisme régnant. Le modèle universitaire napoléonien avait fait *tabula rasa*⁸ et insistait sur une formation scientifique et technique, plus susceptible d'endiguer les mouvements subversifs (Lundgreen 1989 ;

⁸ Les réformes napoléoniennes ont eu pour conséquence de fermer la majorité des universités existantes (Lamy et Gingras 2008).

Charle 2004). L'École Polytechnique de Paris est d'ailleurs née de ce modèle (Daumas 1957), tandis que la littérature scientifique française de l'époque était enrichie de plusieurs travaux sur la mécanique industrielle (Kuhn 1959 ; Cardwell 1967). L'impossibilité du mouvement perpétuel était acceptée, malgré quelques prétentions exprimant ce « désir pathétique d'obtenir quelque chose de rien » (Holton & Brush 2001 : 246). Le polytechnicien Sadi Carnot visait l'examen d'une machine idéale qui, travaillant avec n'importe quel liquide ou vapeur, donnerait le maximum d'effet (Chapitre 4). Or, ses travaux s'inscrivaient dans la tradition des débats entourant la notion de *vis viva* (littéralement *force vive*) sur laquelle ont reposés les développements du concept de *travail*, regroupant ou recoupant plusieurs termes tels que « effet mécanique » (*mechanical effect*), « puissance motrice » (*mechanical power*), « service » (*duty*), « moment d'activité », « effet dynamique », etc. Cette notion de *vis viva* a été apportée par Leibniz, dans un cadre conceptuel favorable aux principes conservatifs, et s'exprime mathématiquement par le produit de la masse et de la vitesse au carré, ou mv^2 (ce qui équivaut, à une constante $\frac{1}{2}$ près, à la formulation moderne de l'énergie cinétique). Le principe de conservation de la *vis viva*, qui a été employée par Lazare Carnot afin d'exprimer l'effet obtenu par une machine, de même que les écrits d'une génération d'ingénieurs théoriciens français plus concernés par l'établissement d'une théorie pratique des machines que par la mécanique rationnelle, ont confirmé ce glissement de la notion de force vers celle de travail, qui l'a finalement remplacée dans ces considérations technologiques (Cardwell 1967 ; Guedj 2006). Ainsi, de considérations principalement technologiques concernant entre autres la performance, la friction et l'ergonomie, l'analyse dynamique traitant des forces conservatives comme la *vis viva* a opéré un changement vers la notion de travail utile et des pertes quantitatives qu'il peut subir (Grattan-Guinness 1984).

3. Le troisième facteur dans l'émergence du concept moderne d'énergie s'inscrit dans le mouvement romantique du tournant du XIX^e siècle. Dans ce contexte d'intenses bouleversements politiques en France et en Allemagne, « la modestie épistémologique était anéantie par le désir de réponses absolues »⁹ (Reill 2008 : 42). La doctrine de la *Naturphilosophie* (littéralement *philosophie de la nature*) a pris racine bien sûr chez Kant, Schelling et Goethe, mais aussi chez Wolff et Leibniz, et visait l'élaboration d'une théorie universelle (*mathesis universalis*), soit une « vision englobante » où l'« esprit », au-delà des contingences des épiphénomènes sensoriels, exprimerait la vraie essence de la réalité. Portée par un universalisme exacerbé, cette « aventure de la raison » souhaitait unifier science et philosophie dans une nouvelle synthèse culturelle, scientifique et esthétique : il s'agissait de récupérer à un autre niveau la vision universelle ayant guidée les philosophies de Platon, Pythagore, Plotin, Descartes et Leibniz ; d'unifier l'esprit et la matière, le vivant et le non vivant, dans un tout uniforme et cohérent ; d'éliminer

⁹ « [...] *epistemological modesty was shattered by the desire for absolute answers.* »

toute contingence ; d'exposer l'histoire de l'univers de ses débuts jusqu'à sa fin (Reill 2008 : 42-3). Kuhn (1959) soutient que cette doctrine a pu jouer le rôle d'un « arrière-plan philosophique », essentiellement en suggérant l'idée de l'unité des forces, car plusieurs pionniers de la conservation de l'énergie, surtout les Allemands, ont commis des écrits présentant des similitudes avec celle-ci. Par exemple, Oersted a longtemps cherché une relation entre électricité et magnétisme en raison de sa conviction philosophique en l'existence d'une force unique, et Colding (l'un des pionniers) était son protégé¹⁰.

L'avènement du concept d'énergie est donc un excellent exemple d'une construction, certes difficile, « longue et tortueuse » (Balian 2013 : 1), mais jouant d'un subtil amalgame entre des considérations métaphysiques, aprioriques, et une attention particulière aux préparations expérimentales, cherchant à quantifier par la mesure. Un dernier exemple, contrastif, est éloquent : pour Mayer (1842, in Sarton 1929), la combinaison de l'indestructibilité et de la convertibilité vient soutenir la thèse que les forces ne s'annihilent pas mais changent de formes, laquelle serait subsumée sous le principe *causa aequat effectum* (la cause égale l'effet) ; tandis que pour Joule (1843, in 1884 : 156), « la quantité de chaleur capable d'augmenter la température d'une livre d'eau d'un degré Fahrenheit est égale à, et peut être convertie en, une force mécanique capable de soulever 8,38 livres d'une hauteur perpendiculaire d'un pied »¹¹.

¹⁰ Bien qu'Helmholtz ait offert le terme « *theoretische Physik* » afin de le distinguer de la *Naturphilosophie*, ce genre d'association ne semblait pas problématique à cette époque puisque la « méthode allemande » se distinguait nettement de celle des Anglais (Seth 2007 : 31). Pour une critique de l'apport de la *Naturphilosophie* au concept de la conservation de l'énergie, voir Brush (1976 : 51-8).

¹¹ « [...] the quantity of heat capable of increasing the temperature of a pound of water by one degree Fahrenheit's scale is equal to, and may be converted into, a mechanical force capable of raising 838 lbs to the perpendicular height of one foot. »

3.2 Thermodynamique

L'histoire de la thermodynamique est particulièrement tortueuse. Elle ne peut manquer d'exposer les incessants allers et retours, essais et erreurs, conjectures et réfutations, contrastant avec les reconstructions contemporaines. Son cas semble particulier pour deux raisons. D'abord, elle semble s'opposer à ce qui était considéré comme les fondements de la physique moderne, notamment les lois de Newton. Ensuite, et partiellement en lien avec la raison précédente, le statut théorique actuel de la thermodynamique fait encore l'objet de multiples restructurations et axiomatisations malgré ses multiples corroborations (Chapitres 4 et 5). À titre d'exemple, on la définit généralement comme la science de l'énergie alors qu'on situe souvent son acte de naissance à une époque où ce concept était loin d'avoir pris consistance, et son nom semble porter à l'erreur puisqu'elle ne traite que des situations d'équilibre ; sont-ce des signes que cette théorie demeure le « morne marais de l'obscurité » (« *Dismal Swamp of Obscurity* », Truesdell 1980 : 6)¹²?

La thermodynamique est essentiellement une théorie visant à caractériser le comportement de corps macroscopiques en termes de quantités physiques, comme la chaleur, le travail, la température et la pression. Ses développements n'ont que très peu été influencés par ceux visant à déterminer la nature ou la structure de la matière. Ses principaux concepts théoriques sont désormais l'énergie et l'entropie. Avec le recul historique, on peut dire que le contenu de la thermodynamique s'est développé à partir de trois ingrédients (Truesdell 1980 ; Uffink 2007).

Le premier est la science de la calorimétrie, étude quantitative des changements dans les corps, comme le changement de volume, associés à l'ajout ou au retrait de chaleur. Elle est donc étroitement liée à la thermométrie, touchant la mesure de la température. Déjà au premier siècle avant J.-C., donc avant même que soient clairement distinguées la chaleur et la température, Héron d'Alexandrie, contestant la thèse aristotélicienne de l'inexistence du vide, utilise un thermoscope, où la dilatation d'une quantité d'air y déplace un volume d'eau. Cette invention fut longtemps oubliée, jusqu'au début du XVII^e siècle, alors que les travaux de Robert Boyle et d'Edme Mariotte (1620-1684), développés plus tard par Laplace et Lavoisier, montrent que le volume d'un gaz dépend de sa température et de sa pression, de sorte que l'indication de température dans un thermoscope dépend de la pression atmosphérique. Au tournant du XVIII^e siècle les thermomètres font leur apparition, suivie un peu plus tard des échelles bien connues de l'Allemand Daniel Gabriel Fahrenheit (1686-1736) et du Suédois

¹² Ce n'est donc pas surprenant que Truesdell (1980) est intitulé son influent livre *The Tragical History of Thermodynamics, 1822-1854* ; un livre où il présente des positions « hérétiques », selon Lervig (1982 : 86), en ce qu'elles s'opposent non seulement à la présentation reçue de l'histoire de la thermodynamique mais aussi de ces concepts théoriques.

Anders Celsius (1701-1744), et celle moins connue du Français René Antoine Ferchault de Réaumur (1683-1757). Joseph Black (1728-1799) fit beaucoup pour la calorimétrie en élaborant le concept de capacité calorifique (C_V), soit celle d'un corps à absorber ou délivrer de la chaleur, et celui de chaleur latente (Λ_T), la chaleur nécessaire pour un changement de phase, comme la fonte de la glace. Ce dernier concept a d'ailleurs inspiré James Watt dans l'invention du condenseur, un dispositif installé aux machines à vapeur dans l'objectif de diminuer leur consommation de charbon nécessaire à la production de vapeur. Ses développements menèrent à cette idée, exprimée dans le langage différentiel et commune « à tous les pionniers de la thermodynamiques » (Truesdell 1980 : 16), que dans chaque processus la chaleur Q est donnée par une fonction linéaire des taux d'accroissement du volume V et de la température T (soit : $\delta Q = C_V dT + \Lambda_T dV$).

Aussi, au milieu du XVII^e siècle est menée une expérience importante, désormais célèbre, visant à comparer la lecture de baromètres placés à différentes altitudes du Puy-de-Dôme et dont les conclusions sont que l'air a un poids et que le vide existe. Peu de temps après, Otto von Guericke (1602-1686) mène une expérience tout aussi célèbre où il fait le vide, à l'aide d'une « pompe à air », à l'intérieur d'une enceinte composée de deux hémisphères de cuivre et attachées à un attelage de huit chevaux qui ne purent les détacher, en raison de la pression atmosphérique et du vide à l'intérieur, cette « force du néant » (« *Energie der Kräfte* »). Cette force est mise à profit avec, bien sûr, la machine à vapeur de type atmosphérique de Thomas Newcomen (1664-1729), améliorée plus tard, tel que mentionné, par Watt. Il a été vu que le fonctionnement de cette machine, mais aussi des machines en général, ont été un facteur primordial dans l'élaboration du concept d'énergie. Ainsi, un autre ingrédient important est l'équivalence du travail et de la chaleur, ces deux moyens de transférer de l'énergie. Plus précisément, on a constaté que lors d'un processus cyclique, la quantité de chaleur absorbée par le système est proportionnelle à la quantité de travail qu'il produit.

Le troisième ingrédient trouve sa source dans les travaux de Carnot (prochaine section) visant à déterminer la quantité de travail maximal pouvant être obtenu d'une quantité de chaleur donnée, caractéristique des machines à vapeur. Cette quantité ne dépend que des températures d'opération du cycle et elle est d'autant plus élevée, c'est donc dire que le rendement est plus grand, si les transferts de chaleur sans augmentation de volume sont limités. L'équivalence entre la chaleur et le travail n'est donc pas symétrique puisque, pour un système donné, la conversion de la chaleur en travail ne peut être complète, il y a irrémédiablement des « pertes ». Il a donc fallu résoudre cet apparent problème de conciliation entre un principe conservatif de l'énergie et celui des pertes d'énergie.

3.2.1 Sadi Carnot

Avec la publication de ses *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance* (1824), Sadi Carnot (1796-1832) peut à juste titre être considéré comme le père de la *thermodynamique classique* (Klein 1974 ; Truesdell 1980 ; Erlichson 1998). L'objet d'étude de Carnot était les « machines à feu », terme plus général que « machines à vapeur », c'est-à-dire des machines thermiques bithermes fonctionnant selon un cycle entre deux réservoirs à températures différentes mais constantes, deux sources de chaleur, à qui l'on doit attribuer « les grands mouvements qui frappent nos regards sur la terre » (Carnot 1824 : 1). L'étude de ces machines était « du plus haut intérêt » car elles semblaient destinées à servir un jour de « moteur universel » (Carnot 1824 : 2). L'objectif principal des *Réflexions* était de répondre à deux questions. D'abord, la puissance motrice de la chaleur, autrement dit le travail¹³, est-elle limitée ou est-elle susceptible d'une « extension indéfinie » ? Ensuite, y a-t-il des agents préférables, autres que la vapeur d'eau, pour développer cette puissance motrice ? Deux *axiomes* soutenaient le raisonnement de Carnot : l'impossibilité d'une machine à mouvement perpétuel (de premier type) et la conservation du calorique (Lervig 1972 ; Darrigol 2003). Pour et par son raisonnement, Carnot a (i) présenté une *hypothèse*, (ii) énoncé un *principe*, (iii) démontré un *théorème* et (iv) en a déduit un *corollaire*¹⁴.

Dans un premier temps, une machine à mouvement perpétuel, au sens où l'entendait Carnot, va au-delà d'un simple « mouvement susceptible de se prolonger indéfiniment après une première impulsion reçue », mais implique aussi la possibilité de « créer de la puissance motrice en quantité illimitée » (1824 : 12, n1), ou « la création de force motrice sans consommation ni de calorique, ni de quelque autre agent que ce soit » (1824 : 11). Or, une telle machine serait « tout à fait contraire aux idées reçues jusqu'à présent, aux lois de la Mécanique et de la saine Physique ; elle est inadmissible »¹⁵ (1824 : 11-12). Dans un second temps, Carnot adopte, à quelques nuances près, la théorie du *calorique*. Carnot n'a fait que prendre assise dans une théorie établie, mais qui commençait à montrer des signes de fatigue alors que des doutes étaient émis à son égard au tournant du XIX^e siècle. D'ailleurs, dès ses *Réflexions* (1824), il affirme que la théorie du calorique ne « paraît pas d'une solidité inébranlable » (50) et que plusieurs « faits d'expérience paraissent inexplicables dans l'état actuel de cette théorie » (20, n1). Des

¹³ « Nous nous servons ici de l'expression puissance motrice pour désigner l'effet utile qu'un moteur est capable de produire. Cet effet peut toujours être assimilé à l'élévation d'un poids à une certaine hauteur [...] » (Carnot 1824 : 4)

¹⁴ Bien que tous s'entendent pour dire que Carnot a eu une grande influence sur le développement de la thermodynamique, près de deux cents ans après la publication de son œuvre, aucun consensus n'existe à propos de la signification réelle de son contenu et plusieurs caractérisations et reformulations sont offertes. (Voir La Mer 1954 ; Kuhn 1955 ; Lervig 1972 ; Klein 1974 ; Truesdell 1980 ; Erlichson 1998 ; Uffink 2001 ; Müller 2003 ; Hertz 2004.)

¹⁵ Le *US Patent Office* a cessé en 1918 de considérer toute demande de brevet à l'intention d'une machine à mouvement perpétuel (Çengel & Boles 2006 : 497).

notes manuscrites postérieures (Carnot 1878) indiquent qu'il a finalement adopté cette interprétation. Néanmoins, en 1824, Carnot adopte la théorie du calorique, faute de mieux.

L'*hypothèse* mentionnée précédemment repose sur une constatation empirique : la production du mouvement ou du travail est toujours accompagnée par le « rétablissement d'équilibre dans le calorique, c'est-à-dire son passage d'un corps où la température est plus ou moins élevée à un autre où elle est plus basse » (1824 : 5-6). Le retour à l'équilibre thermique suivant le transfert de chaleur est alors interprété comme une « chute de calorique », à la base d'une *analogie mécanique* entre le travail provenant de la vapeur et celui provenant d'une chute d'eau où la différence de température est ici assimilée à la différence de hauteur : « [l]a puissance motrice d'une chute d'eau dépend de sa hauteur et de la quantité du liquide ; la puissance motrice de la chaleur dépend aussi de la quantité de calorique employé et de ce qu'on pourrait nommer [...] la *hauteur de sa chute*, c'est-à-dire de la différence de température des corps entre lesquels se fait l'échange du calorique » (Carnot 1824 : 15, italiques originaux).

S'appuyant sur son axiome de la conservation du calorique, Carnot peut alors énoncer son *principe* : « [l]a production de la puissance motrice est donc due, dans les machines à vapeur, non à une consommation réelle du calorique, *mais à son transport d'un corps chaud à un corps froid*, c'est-à-dire à son rétablissement d'équilibre, équilibre supposé rompu par quelque cause que ce soit [...] » (1824 : 6, italiques originaux). Ainsi, une différence de température est à l'origine d'un flux de chaleur, lui-même à l'origine d'un potentiel, d'une « opportunité » selon Brush (1976 : 567) de production de travail. Inversement, il est possible de créer une différence de température à partir d'une consommation de travail. Mais la différence de température est insuffisante : « la chaleur ne peut évidemment être une cause de mouvement qu'en vertu des changements de volume ou de forme qu'elle fait subir au corps [...] » (1824 : 8). Si cette condition n'est pas remplie et que le flux de chaleur est direct, ne produisant donc pas de « changement de volume ou de forme », alors il y aurait une véritable *perte de travail*. Carnot en tire une *conséquence technologique*, à savoir que l'élimination de telles pertes, en raison d'une diffusion de chaleur incontrôlée, serait ce qui constitue la « base fondamentale » de la construction de toute machine thermique. Par ailleurs, afin d'illustrer ses premières conclusions et de les exprimer dans toute leur généralité, Carnot présente ce qu'il est maintenant convenu d'appeler le *cycle de Carnot*, qui décrit des processus isothermiques et adiabatiques (voir Truesdell 1980 : 86-9 pour un point de vue critique).

Carnot utilise ensuite un argument de type *reductio ad absurdum* afin de démontrer son *théorème* à l'effet que les machines *réversibles* présentent un *rendement maximal* : pour une quantité de chaleur (« calorique ») donnée, le travail (« puissance motrice ») est optimal. Soit un corps Θ_+ et un corps Θ_- à des températures différentes respectives T_+ et T_- ($T_+ - T_- = \Delta T$) et deux machines, A et B , la première

présentant un rendement plus faible que la seconde ; si les machines opèrent à partir des mêmes réservoirs a et b , alors la machine A pourrait extraire un travail de ΔT pour faire fonctionner la machine B pouvant rétablir ΔT ; or, comme B présente *ex hypothesi* un meilleur rendement que A , un *surplus* de travail serait obtenu et pourrait être utilisé pour un usage quelconque. Ce procédé est cyclique puisque les deux machines et les deux réservoirs retournent à leurs états initiaux respectifs. Ce serait donc une machine à mouvement perpétuel, en ce sens où il y aurait production de travail sans qu'il n'y d'autre changement, ce qui est impossible ; donc, puisque la conclusion est impossible ou absurde, les prémisses doivent être fausses. Ainsi, le « *maximum de puissance motrice résultant de l'emploi de la vapeur est aussi le maximum de puissance motrice réalisable par quelque moyen que ce soit* » (1824 : 12, italiques originaux). Ce résultat à l'effet que toute machine thermique possède un rendement limité est en soi un développement théorique important et le niveau de généralité souhaité au départ est atteint en raison de l'indépendance de la nature du fluide utilisé (caloporteur), car « tous les corps de la nature peuvent être employés à cet usage » (1824 : 7).

Carnot est alors en mesure de déduire son *corollaire* : la « *puissance motrice de la chaleur est indépendante des agents mis en œuvre pour la réaliser ; sa quantité est fixée uniquement par les températures des corps entre lesquels se fait, en dernier résultat, le transport du calorique* » (1824 : 20, italiques originaux). Il s'ensuit que le rendement η est fonction uniquement des températures des réservoirs chaud (T_+) et froid (T_-). De sorte qu'une relation fonctionnelle f_C , appelé *fonction de Carnot*, est définie telle que (Poincaré 1908 : 44 ; van Wylen, Sonntag & Desrochers 1992 : 189 ; Uffink 2001 : 321) :

$$\eta = W/Q = f_C(T_+, T_-),$$

où W représente le travail effectué par le système et Q la chaleur reçue par le système du réservoir chaud. Cette fonction, qui est supposée universelle, n'a pas été déterminée par Carnot, ni par Kelvin qui s'y est frotté, mais plutôt par Clausius (1850 ; 1854). Elle peut être ramenée à une fonction de la différence de température, tel que $f_C(T_+, T_-) = 1 - T_-/T_+$; elle est ainsi analogue à une « fonction de potentiel » (Lervig 1972 : 231), à l'instar du différentiel de pression dans l'écoulement des fluides.

Il y a ainsi une condition à la *réversibilité* d'un processus cyclique : tout changement de température doit être dû à un changement de volume (soit « la base fondamentale » de la construction des machines thermiques). Ainsi, tout rétablissement d'équilibre sans production de puissance motrice (par le truchement d'un changement de volume) doit être considéré comme une « véritable perte » ; cela équivaut à un rétablissement « inutile » d'équilibre des températures (ou du calorique). Or, s'il y a un tel rétablissement inutile, alors le processus n'est pas réversible et il est donc *irréversible*¹⁶. Cela signifie

¹⁶ La notion d'irréversibilité est riche et problématique (voir Chapitre 4).

que la quantité de travail qui *aurait pu être* utilisée est *irré récupérable*. Supposons, par exemple, et comme précédemment, qu'un système soit constitué de deux corps à des températures différentes : si le cycle est réversible, il est possible de rétablir les conditions initiales du système, c'est-à-dire de rétablir la température initiale des deux corps, donc la différence de température. Une certaine quantité de travail $W_{\text{rév}}$ demeure donc disponible. Par contre, si le cycle est irréversible, il n'est pas possible de rétablir les conditions initiales du système et la quantité de travail disponible est $W_{\text{irrév}} < W_{\text{rév}}$, de sorte que la perte de travail $W_p = W_{\text{rév}} - W_{\text{irrév}}$. Pour avoir accès à une quantité de travail équivalente à $W_{\text{rév}}$ afin de compenser la quantité W_p , et donc de recouvrer les conditions initiales, il faudra ouvrir les limites du système. Or, le nouveau système sera soumis aux mêmes contraintes et ultimement il faudra dépasser les limites de l'univers, ce qui est évidemment impossible. Le cycle de Carnot présente donc, pour une différence de température donnée, un rendement maximal. Le problème est qu'il est impossible nous de le construire.

En somme, l'apport conceptuel de Carnot à la thermodynamique a souvent reçu une appréciation mitigée en raison de son adoption de la théorie calorique et plus spécifiquement de sa conservation, qui demeure incompatible avec le premier principe : il aurait ainsi obtenu un résultat théorique valide en utilisant une théorie de la chaleur erronée. Néanmoins, Carnot peut être considéré à juste titre comme le père de la thermodynamique classique¹⁷. En effet, en insistant sur l'importance d'une différence de température à l'obtention d'un travail, il a ouvert la voie au principe d'interconvertibilité de la chaleur en travail, malgré son adhésion à celui de la conservation du calorique, hypothèse qu'il a d'ailleurs par la suite récusée. De plus, la conceptualisation d'un cycle idéal présentant un rendement maximal lui a permis de rendre compte des pertes de rendement résultant de ce qui est maintenant convenu d'appeler les irréversibilités, ce qui a été essentiel à l'heuristique de l'élaboration conceptuelle de l'entropie comme concept complémentaire à celui d'énergie.

Malheureusement, les travaux de Carnot ont été longtemps ignorés (Cardwell 1971 ; Redondi 1976 ; Fridberg 1978 ; Hertz 2004). L'éditeur des *Réflexions*, Bachelier, semble n'avoir fait aucun effort ni publicité pour faire connaître ce traité, tandis qu'on ne retrouve aucune mention dans le catalogue de l'École Polytechnique, ni dans les *Annales de Chimie et de Physique* ou les *Annales des mines*, pas plus que dans les *Bulletins des Sciences technologiques*. Émile Clapeyron (1799-1864) fait figure d'exception, lui qui revenait d'un séjour à Saint-Petersbourg où il avait enseigné aux ingénieurs russes, découvrit à l'Association Polytechnique un exemplaire des *Réflexions*, dont il fit un résumé dans un mémoire publié

¹⁷ Cependant, si l'on définit la thermodynamique comme la science de l'énergie et de l'entropie, en toute rigueur Carnot ne peut être le « père de la thermodynamique ». Il serait alors plus juste de dire qu'il en serait le « grand-père » en ce qu'il a jeté les bases de son élaboration en établissant des résultats importants quant aux processus impliquant des échanges de chaleur et de travail.

dans le *Journal de l'École Polytechnique* en 1834. Quelques ingénieurs français ont par la suite appliqué les idées de Carnot à des procédés technologiques, mais l'influence reste marginale ; par exemple, Boucherot présenta un « pyroaéromoteur » en 1835 dans les *Comptes rendus de l'Académie des sciences* et Bresson présenta une « machine à air chaud » en 1836 aussi dans les *Comptes rendus*. Lors d'un stage au laboratoire de Henri-Victor Régnault, William Thomson trouva le mémoire de Clapeyron, qui retint son attention. Il remonta à la source et ses articles subséquents sur la théorie de chaleur contribuèrent à la notoriété des *Réflexions*.

3.2.2 William Thomson, alias Baron Kelvin

Après avoir gradué de l'Université de Cambridge à titre de *Second Wrangler*¹⁸ sous la direction de William Hopkins (1793-1866), William Thomson (1824-1907), ayant déjà publié plusieurs articles en physique mathématique principalement basés sur la méthode développée par Fourier, entreprend en 1845 un stage dans le laboratoire de Henri-Victor Régnault (1810-1878), qui travaille alors sur l'élasticité des fluides et des gaz et sur leurs chaleurs spécifiques. Après une première rencontre lors d'une conférence à Oxford, il collabore, en 1849, avec Joule, qui avait mesuré deux fois l'équivalent calorique du travail par deux expériences calorimétriques. En reconnaissance de sa contribution à la science – 661 articles et 69 brevets (Cropper 2001 : 79) – il est élevé, en 1892, à un rang nobiliaire britannique en tant que *Baron Kelvin of Largs* (Daintith 2009 : 406).

Dans son article « Un compte-rendu de la théorie de Carnot sur la puissance motrice de la chaleur ; avec des résultats numériques déduits des expériences de Régnault sur la vapeur »¹⁹ (1849, in 1882), William Thomson (1824-1907) entreprend une formalisation analytique de la théorie de Carnot afin d'expliquer l'origine de la puissance motrice de la chaleur. Ce texte, comme plusieurs autres de cette période, est traversé par une ambiguïté entre les termes de *travail* et d'*effet mécanique*, le premier étant généralement appliqué à l'action directe d'une force et le second à l'établissement d'un bilan en terme de travail. Il s'appuie sur le raisonnement de Carnot à savoir que l'effet mécanique est obtenu du transfert (« *transference* »²⁰) de la chaleur d'un corps chaud à un corps froid : lorsqu'un corps effectue un cycle et retourne à son état initial et qu'il a absorbé durant ce cycle une certaine quantité de chaleur, il doit avoir émis la même quantité de chaleur. La question, qui est tout autant pratique que théorique, se

¹⁸ À l'université de Cambridge, un *Wrangler* est un étudiant qui a obtenu les meilleurs résultats scolaires en troisième année de mathématiques (cam.ac.uk 2011).

¹⁹ « *An Account of Carnot's Theory of the Motive Power of Heat; with Numerical Results deduced from Regnault's Experiments on Steam.* »

²⁰ « Transport » chez Carnot (1824 : 6).

pose alors, comme chez Carnot, de la quantité de chaleur ou de la différence de température (« *thermal agency* ») nécessaire à l'obtention d'un effet mécanique.

La stratégie qu'emploie Thomson pour répondre à cette question repose sur la notion de réversibilité et passe par l'analyse d'une « machine thermodynamique parfaite » – la première fois où apparaît le mot « *thermodynamique* » – où, d'une part, aucune chaleur ne serait « gaspillée »²¹ et où, d'autre part, la chaleur fournie pour produire un certain effet mécanique peut être restituée par l'action d'un même effet mécanique. Or, demande Thomson (1849, in 1882 : 118, n1), qu'en est-il de cette « perte » si rien ne peut être annihilé dans les « opérations de la nature », si « aucune énergie ne peut être détruite »²² ? Il affirme que malgré l'ébauche d'une solution dans les travaux de Joule, l'état actuel de la science ne fournit pas de réponse à cette question. Or, le terme « énergie » mentionné ici n'est pas associé à un concept de la physique mais apparaît plutôt comme « un *a priori* métaphysique auquel les lois de la physique doivent se soumettre » (Guedj 2006 : 40). Ce n'est plus tard que Thomson proposera véritablement un concept physique précis pour la notion d'énergie.

Dans son article « De la théorie dynamique de la chaleur ; avec des résultats numériques déduits de l'équivalent d'unité thermique de M. Joule et des observations de M. Régnauld sur la vapeur »²³ (1851, in 1882), Thomson énonce deux propositions sur lesquelles repose la « théorie de la puissance motrice de la chaleur » et qui peuvent aisément être assimilées aux deux premiers principes de la thermodynamique classique. La première proposition (attribuée cette fois à Joule et non à Mayer) stipule justement l'équivalence entre travail mécanique et chaleur.

PROPOSITION I. Lorsque des quantités égales d'effet mécanique sont produites par un quelconque moyen à partir d'une source de chaleur ou perdues en effets thermiques, des quantités égales de chaleur sont éliminées ou générées.²⁴

Elle s'appuie sur le raisonnement suivant : lorsque le système effectue un cycle (réversible), l'effet mécanique qui correspond au travail doit être équivalent à la quantité de chaleur absorbée ou émise par le système. Il y a donc un *bilan* qui est effectué entre les effets mesurés de la chaleur et du travail.

La seconde proposition (attribuée à Carnot et à Clausius) énonce le cas limite que présentent les machines réversibles.

²¹ Ce qui correspond selon Carnot à un rétablissement « inutile » des températures et à ce qui est présenté comme une « perte » d'effet mécanique.

²² « [...] *no energy can be destroyed* ».

²³ « *On the Dynamical Theory of Heat ; with Numerical Results deduced from Mr Joule equivalent of a Thermal Unit, and M. Regnauld's Observations on Steam.* »

²⁴ *PROP I. When equal quantities of mechanical effect are produced bay any means whatever from purely thermal sources, or lost in purely thermal effects, equal quantities of heat are put out of existence or are generated.* (1851, in Thomson 1882 : 178)

PROPOSITION II. Si une machine est telle que, lorsque son fonctionnement est renversé, ses éléments physiques et mécaniques sont tous renversés, alors elle produit autant d'effet mécanique qu'il est possible d'en produire d'une machine thermodynamique, pour des températures du foyer et du réfrigérant données, à partir d'une quantité de chaleur donnée.²⁵

Cette proposition est basée sur un axiome stipulant l'impossibilité des machines à mouvement perpétuel de second type, à savoir qu'il est impossible d'obtenir un travail en refroidissant un corps à une température inférieure à l'environnement, et sur le raisonnement de type *reductio ad absurdum* de Carnot sur les machines thermodynamiques réversibles, définissant une *machine thermodynamique parfaite* selon un cycle *réversible*. Cet axiome est équivalent, selon Thomson, à celui de Clausius en ce sens qu'il est impossible, pour une machine autonome (« *self-acting machine* ») et sans « agent externe », de transmettre de la chaleur d'un corps chaud à un corps froid. La notion d'irréversibilité est ici à l'œuvre, car on y définit deux processus dont l'un est « spontané » et l'autre ne l'est pas : aucun travail n'est nécessaire pour transmettre de la chaleur d'un corps chaud à un corps froid, ce qui correspond à un processus spontané, tandis que pour transmettre de la chaleur d'un corps froid à un corps chaud, un travail est nécessaire²⁶ ; de plus, comme la conversion de la chaleur en travail n'est jamais totale ou parfaite (le coefficient de Carnot est inférieur à l'unité), la quantité de chaleur transmise au corps froid (moins chaud) ne peut plus être convertie en travail ; elle est « irrémédiablement perdue pour l'homme, et donc 'gaspillée', quoiqu'elle ne soit pas *annihilée* »²⁷ (Thomson 1851, in 1882 : 189). Thomson est donc passé, grâce à Carnot et à Joule, de la *conservation de l'énergie* à la *perte de travail*. Ce qui est une version quelque peu simplifiée du second principe de la thermodynamique et annonce la notion d'entropie.

Ainsi, le problème des pertes trouve sa réponse dans l'appréhension de la dissipation de l'énergie dans le contexte de sa conservation. C'est pourquoi le second principe de la thermodynamique prend appui sur le premier ; ou, de manière plus colorée, c'est pourquoi « il n'a pu être baptisé avant d'avoir donné naissance à son frère » (Brush 1976 : 569). Mais Thomson n'a jamais utilisé la notion d'entropie telle que définie par Clausius. Cela étant, dans son article « De la tendance universelle dans la nature vers la dissipation de l'énergie mécanique »²⁸ (1852, in 1882), il présente un axiome, sur lequel reposerait sa proposition II et qui est parfois considéré comme une formulation valable du second

²⁵ PROP II. If an engine be such that, when it is worked backwards, the physical and mechanical agencies in every part of its motions are all reversed, it produces as much mechanical effect as can be produced by any thermo-dynamic engine, with the same temperatures of source and refrigerator, from a given quantity of heat. (1851, in Thomson 1882 : 178)

²⁶ Si les tasses de café se refroidissent d'elles-mêmes, les réfrigérateurs par contre doivent être branchés pour fonctionner.

²⁷ « [...] irrecoverably lost to man, and therefore 'wasted', although not annihilated. »

²⁸ On a Universal Tendency in Nature to the Dissipation of Mechanical Energy.

principe (511, italiques originaux) : « *il est impossible, par le truchement d'agent matériel inanimé, d'obtenir un effet mécanique à partir d'un corps qui serait refroidi à une température inférieure à celle des objets environnants* »²⁹. Une illustration commune de cette formulation concerne l'impossibilité, intuitive à tout le moins, pour un navire d'avancer en tirant profit du fantastique réservoir d'énergie que constitue l'océan et ne laissant derrière lui qu'un sillage d'eau plus froide.

Thomson présente plus loin ce qui pour certains (par ex. Lindley 2004 : 110), après quelques ajustements, pourraient être identifié à l'entropie. Il affirme qu'il y a « à l'heure actuelle » (*sic*) une tendance universelle dans le monde matériel à la dissipation de l'énergie mécanique. Il en conclut que la terre deviendra tôt ou tard « inhabitable pour l'homme » (1852, in 1882 : 514), et que cela est « mécaniquement inévitable » (1855, in 1884 : 37). Ces considérations géologiques ont mené à la détermination de l'âge de la terre, participant ainsi aux intenses débats concernant la théorie darwinienne de l'évolution. Cependant, cette thèse sur la dissipation de l'énergie, bien qu'importante en regard du second principe, n'a apparemment joué aucun rôle dans l'élaboration de ce qu'il a appelé « l'expression mathématique de la seconde loi fondamentale de la théorie dynamique de la chaleur » (Thomson 1854, in 1882 : 237), qui répond selon lui au critère de réversibilité, au sens de récupérabilité (Chapitre 4), et s'exprime comme la sommation discrétisée des chaleurs transférées Q_i divisées par la température locale T_i : $\sum Q_i/T_i = 0$. Cela est très semblable à l'expression donnée par Clausius, pour les cycles fermés réversibles, quelques mois plus tard.

3.2.3 Rudolf Clausius

Né d'un père ministre de l'église de Köslin et conciliateur au Gouvernement Royal de Prusse, Rudolf Julius Emmanuel Clausius (1822-1888) a étudié les mathématiques et la physique à l'Université de Berlin où il devint *Privatdozent*, avant d'enseigner à l'École royale d'ingénieurs et d'artillerie à Berlin, à l'Université de Zurich, de Würzburg, puis enfin à celle de Bonn. Comme Thomson, avec lequel il échangeait d'ailleurs régulièrement ses publications, il tenta d'interpréter théoriquement la couleur bleue du ciel mais il échoua où Thomson réussit en faisant l'hypothèse de la diffusion.

Il a néanmoins établi la fonction d'état définissant le concept d'entropie, qu'il a lui-même baptisé afin d'insister sur la notion de « transformation » et sur son lien avec le concept d'énergie. Selon lui, la thermodynamique devait se baser sur ces deux concepts que sont l'énergie et l'entropie.

²⁹ « *It is impossible, by means of inanimate material agency, to derive mechanical effect from any portion of matter by cooling it below the temperature of the coldest of the surrounding objects* ».

Dans l'article « Sur la force motrice de la chaleur et les lois qui s'en déduisent pour la théorie même de la chaleur »³⁰ (1850, in 1868 ; aussi in Magie 1899), Clausius réconcilie les travaux de Carnot avec ceux de Joule et présente ce qui paraît « suffisant à l'élaboration de la thermodynamique classique » (Truesdell 1980 : 196). Dans un premier temps, il adhère à l'une des thèses de Carnot à l'effet que, lorsqu'un travail est effectué par la chaleur et qu'aucun changement ne se produit dans l'état du corps (caloporteur), une certaine quantité de chaleur passe d'un corps chaud à un corps froid ; mais il récuse celle de l'indestructibilité de la chaleur et adopte plutôt l'hypothèse fondamentale de l'interconvertibilité du travail et de la chaleur. Il passe ainsi outre les prétendues difficultés insurmontables de Thomson et soutient (avec Joule) que la *production* d'un travail implique la *consommation* d'une quantité proportionnelle de chaleur. Toutefois, dans les échanges d'un système avec son environnement (« travaux extérieurs »), il est difficile d'en dresser un bilan complet puisqu'il peut y avoir des modifications internes (« travaux intérieurs »). Dans un deuxième temps, à partir du bilan déterminé lors d'un cycle de Carnot entre la chaleur et le travail, Clausius est en mesure de donner une expression analytique (équation différentielle) du principe d'interconvertibilité et de définir une fonction d'état U (appelée par Thomson « *energy* » et aujourd'hui l'*énergie interne*), soit une fonction arbitraire du volume et de la température, qui exprime la somme de la « chaleur libre » (énergie cinétique des molécules) et de la « chaleur consommée lors d'un travail » (énergie potentielle des forces intermoléculaires). La distinction des grandeurs intérieures et extérieures permet ainsi d'affirmer la conservation de la chaleur, tandis que la conception dynamique de la chaleur soutient l'inexistence de la chaleur latente. La chaleur et le travail (extérieur) ne sont donc plus exprimés par des fonctions d'état. Avec l'hypothèse des gaz parfaits, Clausius présume (correctement) que U n'est fonction que de la température, ce qui lui permet de calculer la fonction de Carnot : $F(T) = 1/T$, où T est la température absolue.

Dans son article « Sur une forme modifiée du second principe de la théorie mécanique de la chaleur »³¹ (1854, in 1868) Clausius va au-delà du principe de la conservation de l'énergie (premier principe) et s'aventure dans l'analyse de celui de la dégradation de l'énergie (second principe). Il statue d'abord sur l'interconvertibilité et la proportionnalité du travail et de chaleur, qui s'appuient sur la distinction entre les notions de travail intérieur et travail extérieur, lesquelles reçoivent l'expression analytique du mémoire de 1850. Par la suite, il soutient que le principe de Carnot, mis « en harmonie » avec le premier principe, exprime une relation entre deux types de transformations : une

³⁰ « Ueber die bewegende Kräfte der Wärme und die Gesetze, welche sich daraus für die Wärmelehre selbst ableiten lassen. »

³¹ « Ueber eine veränderte Form des zweiten Hauptsatzes der mechanischen Wärmetheorie ». Selon Darrigol (2001 : 3), il s'agit de « l'un des plus étranges mémoires de toute l'histoire de la physique ».

transformation d'une quantité de chaleur à une température élevée à une température plus basse (que l'on qualifie souvent de « spontanée » ou « naturelle ») et une transformation de chaleur en travail (parfois appelée « forcée » ou « non naturelle »). Son objectif est alors de donner une expression analytique de ces transformations pour un processus cyclique réversible dans lequel celles-ci semblent se compenser mutuellement par des « valeurs d'équivalence ». Ce faisant, Clausius entreprend « l'un des raisonnements les plus prometteurs de l'histoire des sciences » (Cropper 2001 : 98). Il propose alors ce qu'il présente comme un axiome, censé remplacer le théorème de Carnot et qui est généralement considéré comme une formulation valable du second principe : « *[i]l ne peut jamais passer de chaleur d'un corps froid à un corps plus chaud, à moins qu'il ne se présente en même temps une autre modification qui dépende de la première* » (138, italiques originaux). Évidemment, cette formulation a tout ce qu'il faut pour passer un long séjour dans les limbes des interprétations interminables. Et Clausius semble s'en rendre compte car il présente une autre formulation : « *[I]a chaleur ne peut passer d'elle-même d'un corps froid à un corps plus chaud* » (138, n1, italiques originaux) ; aussitôt ajoute-il que les mots « d'elle-même » ont besoin d'éclaircissement, mais les traduit par « sans compensation », ce qui n'est guère plus éclairant.

L'objectif général de (1854) peut être décrit comme une tentative de donner une réponse au problème de la *récupérabilité*, c'est-à-dire le rétablissement des conditions initiales, qui trouve sa contrepartie mathématique dans l'*intégrabilité* qui en est une condition nécessaire et suffisante (Dias, Pinto & Cientiano 1994 : 154). C'est essentiellement à cette tâche que se voue le traitement mathématique de Clausius, qui est évidemment trop long pour être présenté ici dans le détail ; mais en voici les grandes lignes. D'abord, dans un *cycle fermé réversible*, les deux types de transformations (susmentionnés) sont équivalents et se compensent mutuellement. Dans un tel processus impliquant des échanges (infinitésimaux) dQ à une température T , le *critère de compensation* est exprimé par

$$\int f(T) dQ = \int dN = 0,$$

où $f(T)$ est une fonction universelle de la température, soit la fonction indéfinie de Carnot, et N , qui sera remplacé plus tard par le symbole consacré S , représente une fonction générale d'état.

Un argument mathématique permet à Clausius de définir $f(T) = 1/T$ où T est la température absolue. Voici brièvement de quoi il retourne (pour plus de détails, voir Clausius 1879 ; Truesdell 1980 ; Copper 1986 et 2001 ; Müller 2007). Clausius utilise d'abord une équation calorimétrique relativement connue à l'époque, une équation établissant un rapport linéaire entre la chaleur transférée et les taux d'accroissement du volume et de la température³². Il est ensuite en mesure de dériver une équation obtenue à partir de cette équation calorimétrique, du critère de compensation susmentionné

³² Ce qui lui permet d'écrire : $dN = f(T) dQ = f(T) M dV + f(T) N dT$, où M et N sont deux constantes.

et d'un résultat de son (1850)³³. En appliquant cette dernière équation à un gaz parfait et après quelques manipulations élémentaires, on obtient : $f(a + t) = -f$, où a est une constante et t est la température mesurée. En prenant a comme étant égale à 273, on peut remplacer $(a + t)$ par la température absolue T , en sorte que $f(T) = 1/T$. Ainsi, pour un *cycle fermé réversible*, l'« expression analytique du second principe de la théorie mécanique de la chaleur » (1854, in 1868 : 153) est :

$$dN = dQ/T \text{ [cycle fermé réversible]}.$$

Cette différentielle exacte (donc intégrable) est nulle et correspond ainsi à la quantité conservative associée à la récupérabilité des conditions initiales (au lieu de δW chez Carnot). En revanche, pour un *cycle fermé non réversible*, Clausius (1854, in 1868 : 157) ne donne pas d'expression mathématique mais affirme que la somme des transformations ne peut être que positive, ce qui est généralement interprété (et ce le sera plus tard par Clausius lui-même) par :

$$dS > dQ/T \text{ [cycle fermé non réversible]}.$$

Ainsi, en 1854, Clausius a défini une fonction d'état S , qu'il n'a pas nommée et qui exprime, pour l'heure de manière plutôt absconse, les « valeurs d'équivalence » lors de cycles fermés réversibles ou non réversibles. De plus, il parle encore du « second principe de la théorie mécanique de la chaleur » et propose « principe de l'équivalence des transformations ».

Si l'acte de naissance de la fonction d'état représentant l'entropie est venu dans le (1854) de Clausius, son baptême a été célébré dans son article « Sur diverses formes des équations fondamentales de la théorie mécanique de la chaleur, qui sont commodes dans l'application »³⁴ (1865 ; 1868). Clausius y unifie la valeur de transformation d'un corps (somme algébrique des transformations qui se produisent dans un cycle) et celle de son niveau de désagrégation (action de la chaleur) par la célèbre grandeur qu'il nomme « *entropie* ». Il s'agit en somme de caractériser la fonction d'état intégrable associée à la récupérabilité des conditions initiales.

Si, comme elle l'exprime, l'intégrale $\int dQ/T$ doit être nulle chaque fois que le corps, en partant d'un état initial et en parcourant une série quelconque d'autres états, revient de nouveau à celui-là, l'expression dQ/T qui se trouve sous le signe d'intégration doit être la différentielle totale d'une quantité qui ne dépend que de l'état actuel du corps, et non de la voie par laquelle il y arrive. (1868 : 408-9)

Cette intégrale est alors désignée par S (précédemment \mathcal{M}), pour aucune raison apparente (Battino, Strong & Wood 1997 ; Howard 2001). Elle correspond au « contenu de transformation d'un corps »,

³³ Soit $f(T)M = -A f(T) [\partial P / \partial T]_V$, où A est une constante.

³⁴ « *Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie* ».

c'est-à-dire à sa capacité de subir des transformations « spontanées » : plus grande est la valeur de l'entropie, plus faible est la capacité du corps à se transformer et plus élevé est son niveau de désagrégation. Pour le nom « entropie », Clausius (1865 : 890 ; 1868 : 411) s'explique :

Si l'on cherche un nom caractéristique pour S , on pourrait lui donner celui de *contenu de transformation*, de même qu'on a nommé U le *contenu de chaleur et d'œuvre*. Mais je pense qu'il est préférable d'emprunter aux langues anciennes pour les quantités scientifiques aussi importantes, afin que ce nom puisse être utilisé sans modification pour toutes les langues vivantes, je suggère donc que le terme S soit appelé l'*entropie* du corps du mot grec $\eta\tau\rho\omicron\pi\eta$ pour « transformation » [ou « forme »]. C'est à dessein que j'ai formé ce mot *entropie*, de manière qu'il se rapproche le plus possible du mot *énergie* ; car ses deux quantités ont une telle ressemblance dans leur signification physique qu'une certaine analogie de dénomination m'a paru utile.³⁵

Il conclut son article par une formulation cosmologique des deux principes fondamentaux de la thermodynamique :

L'énergie du monde est constante.

L'entropie du monde tend vers un maximum.³⁶

Clausius signe donc l'acte de naissance du concept d'entropie en le nommant et en le formulant mathématiquement. Ce faisant, il donne la formulation « classique » des deux premiers principes de la thermodynamique. (Pour la suite, cependant, le second principe de la thermodynamique, exprimant l'impossibilité d'une conversion complète de chaleur en travail, sera distingué du principe entropique, stipulant l'augmentation de l'entropie.) Le vocable « entropie » ne s'est toutefois pas imposé immédiatement dans la littérature anglaise, Peter Guthrie Tait, qui a été suivi par Maxwell, lui attribuant la signification d'« énergie disponible moyenne » (Klein 1983). Les Français et les Allemands ont toutefois été plus réceptifs et à la fin du XIX^e siècle le concept d'entropie faisait partie de « l'arsenal de la physique » (Darrigol 2003).

³⁵ *Sucht man für S einen bezeichnenden Namen, so könnte man, ähnlich wie von der Gröfse U gesagt ist, sie sey der Wärme- und Werkinhalt des Körpers. Da ich es aber für besser halte, die Namen derartiger für die Wissenschaft wichtiger Größen aus den alten Sprachen zu entnehmen, damit sie unverändert in allen neuen Sprachen angewandt werden können, so schlage ich vor, die Gröfse S nach dem griechischen Worte $\eta\tau\rho\omicron\pi\eta$ die Verwandlung, die Entropie des Körpers zu nennen. Das Wort Entropie habe ich absiebtlich dem Worte Energie möglichst lähnlich gebildet, denn die beiden Gröfßen, welche durch diese Worte benannt werden sollen, sind ihren physikalischen Bedeutungen nach einander so nahe verwandt, dafs eine gewisse Gleichartigkeit in der Benennung mir zweckmäßig zu sein scheint.* (Traduction française in Clausius 1868.)

³⁶ « *Die Energie der Welt ist constant. Die Entropie der Welt strebt einem Maximum zu.* » (Clausius 1865 : 400 ; 1868 : 420)

3.2.4 Aparté : Rankine et Poincaré

William John MacQuorn Rankine (1820-1872) peut être considéré comme « l'un des leaders dans le développement de la thermodynamique » (Brush 2003 : 19)³⁷, bien que ses contributions sont parfois perçues comme « éphémères » (Hutchison 1981 : 10), avec un style « obscur » (Lindley 2004 : 211). À l'instar de Thomson, l'un des rares à l'avoir lu de son vivant, il croyait fermement à la primauté du concept d'énergie selon une approche mécaniste et à la thèse dynamique de la chaleur. Ses considérations phénoménologiques étaient doublées d'un modèle hypothétique compliqué du comportement de molécules entourées de tourbillons d'« atmosphère élastique ». Malgré ces idées spéculatives, symptôme d'un esprit néanmoins créatif, il a su introduire dès 1854 une fonction très semblable à celle de l'entropie de Clausius, qu'il a malheureusement appelé « fonction thermodynamique ». Mais il n'a jamais parlé de l'augmentation de l'entropie, concept auquel il s'est d'ailleurs opposé, se limitant, ce qui n'est tout de même pas rien, à définir l'équivalent de la fonction d'état pour le cas limite idéalisé d'un cycle réversible.

Henri Poincaré (1854-1912) fut bien sûr un scientifique extraordinaire, peut-être le « dernier grand savant universel » (Vuillemin, in Poincaré 1968 : 5). Nul besoin de revenir ici sur ses grandes réalisations. En ce qui nous concerne plus particulièrement, il fut un commentateur perspicace des premiers développements de la thermodynamique et de la mécanique statistique. Dans son ouvrage *Thermodynamique* (1892), des leçons professées pendant le semestre 1888-89 à la Faculté des sciences de Paris, il évoque la possibilité « d'élever sur la Thermodynamique seule l'édifice tout entier de la Physique mathématique » (v)³⁸. Il a su par ailleurs identifier correctement le mouvement brownien en 1904 mais aussi en déduire qu'à cette échelle, le second principe, le « principe de Carnot », était violé (Müller 2007 : 274).

3.2.5 Entropie de non-équilibre

Les limites de la thermodynamique classique, ou thermostatique, ont vite été constatées. Se limitant pour l'essentiel aux phénomènes réversibles et à l'équilibre, en bref à des idéalizations, le besoin d'une « extension » s'est fait rapidement sentir. Le concept d'entropie ne fait pas exception puisque la fonction d'état qui l'encadre n'est également que définie pour les processus réversibles et à l'équilibre,

³⁷ Pourtant, le livre de plus de 300 pages de Müller (2007), *A History of Thermodynamics*, ne mentionne pas une seule fois son nom.

³⁸ Ni Uffink (2001) ni Müller (2007) ne mentionne cet ouvrage. Mais Magie (1899) cite « Poincaré *Thermodynamique* ».

ou quasi-statiques. Il faut donc une thermodynamique du non-équilibre, ou si l'on veut, une thermodynamique à proprement dit. Bien que fort difficile comme théorisation, il y a fort à parier que la stratégie explicative ne soit pas radicalement différente. De fait, malgré ce besoin de généralisation, le concept d'entropie est tout de même considéré comme valide avec l'inégalité de Clausius-Duhem. Puisque cette entreprise est relativement récente, elle ne sera pas présentée dans le détail (voir Yourgrau, van der Merwe & Raw 1966 ; Prigogine 1968 ; Le Bellac, Mortessagne & Batrouni 2004 ; Müller 2010).

Avant même qu'il y ait une thermodynamique du non-équilibre, et même avant la thermodynamique tout court, il y avait des « équations phénoménologiques », des équations gouvernant les flux de diverses quantités physiques comme l'énergie. L'une de ses équations a été proposée par Jean Baptiste Joseph Fourier (1768-1830), qui a écrit *Théorie de la chaleur* (1822), un ouvrage très important pour la physique mathématique de l'époque et qui a eu une grande influence par la suite, notamment grâce à sa méthode d'analyse harmonique, l'analyse par les « séries de Fourier ». D'ailleurs, en introduction, il fait une remarque semblable à celle que fera Carnot deux ans plus tard sur la nécessité d'étudier ce phénomène universel de la chaleur, dont l'action « est toujours présente, elle pénètre tous les corps et les espaces, elle influe sur les procédés des arts [techniques], et concourt à tous les phénomènes de l'univers » (Fourier 1822 : 1-2). La loi de Fourier exprime le flux de chaleur, son taux de transfert par unité de surface perpendiculaire à la direction de ce transfert, et proportionnel au gradient de température. D'autres équations de ce type sont celle d'Ohm, qui exprime le courant électrique par unité de surface, ou densité de courant, perpendiculaire à la direction de ce transfert, et proportionnel au gradient du potentiel électrique, ou encore celles de Navier-Stokes décrivant le mouvement des fluides. Ces équations ont reçu une formalisation les généralisant tel que

$$\mathcal{J} = LX,$$

où X représente un gradient ou une affinité, appelé aussi « force généralisée », \mathcal{J} un flux ou un taux de réaction, appelé aussi « flux généralisé », et L est un coefficient indépendant appelé « coefficient phénoménologique ». On peut alors considérer que la force généralisée X « cause » le flux \mathcal{J} (Yourgrau, van der Merwe & Raw 1966 ; Prigogine 1968). La production d'entropie, σ , pour un processus de non-équilibre peut alors s'exprimer comme

$$\sigma = d_e S/dt = \sum \mathcal{J}_i X_i.$$

Lars Onsager (1903-1976) a obtenu le prix Nobel de chimie en 1968 pour son « théorème de réciprocité » à l'effet que les coefficients L sont symétriques, comme quoi une force peut produire un gradient, et vice-versa.

Ces outils conceptuels ont été utilisés dans l'une des premières tentatives de théoriser le non-équilibre ayant été apporté avec la thermodynamique des processus irréversibles. L'intérêt d'une telle théorie est primordial puisque la plupart des processus physico-chimiques comme la diffusion, la conduction de la chaleur, les réactions chimiques sont considérés comme des processus irréversibles. S'appuyant sur les travaux de Fourier, de Navier et de Stokes, mais aussi d'Adolf Fick (1829-1901), Carl Henry Eckart (1902-1973), aussi connu pour ses travaux en mécanique quantique, a su apporter une contribution majeure au domaine avec deux articles parus en 1940 dans *Physical Review*. On peut affirmer que l'objectif de la thermodynamique de non-équilibre est la détermination de champs pour un fluide donné, celui de la densité de masse, ceux de la vitesse dans les trois directions spatiales et celui de la température. Et elle s'appuie sur deux hypothèses fondamentales. La première est l'*hypothèse du continuum*, comme quoi les quantités physiques comme la masse ou l'énergie sont considérées uniformément réparties pour un petit élément de volume et la matière est infiniment divisible. Les champs peuvent être calculés à partir de quelques équations fondamentales de la mécanique dite des « milieux continus », comme la deuxième loi de Newton et le premier principe de la thermodynamique. Mais pour bien être mesure de calculer ces champs avec ces équations, autrement dit pour « fermer » le système d'équations, il faut établir des relations entre certaines variables. Ces relations sont établies par une équation, établie par Gibbs, donnant le taux de variation de l'entropie dans le temps, dont les termes sont déterminés *comme si* le système était en état d'équilibre. Cette seconde hypothèse fondamentale est l'*hypothèse de l'état local associé* : bien que le système soit en mouvement, et donc dans un état de non-équilibre, chaque unité de volume élémentaire peut être considéré comme approximativement en équilibre du point de vue thermodynamique.

Une autre tentative, développée dans les années soixante, est celle de la *thermodynamique rationnelle*, qui a d'abord été développée par Coleman & Noll (1963), puis défendue par Truesdell, et qui présente plusieurs points communs avec la thermodynamique des processus irréversibles. Elle *postule* une inégalité associée à l'entropie, l'inégalité de Clausius-Duhem, et *dérive* l'équation de Gibbs pour la partie de l'équilibre de l'entropie spécifique. L'inégalité de Clausius-Duhem, fondamentale en thermodynamique de non-équilibre, s'appuie sur l'*existence* de l'entropie mais aussi sur son augmentation dans le temps, d'où l'inégalité. Bien que cette théorie n'a pas à s'appuyer l'hypothèse de l'état local associé, donc sur un principe d'équilibre local, elle doit faire appel à une forme spécifique de l'entropie sans réelle justification. Mais la thermodynamique des processus irréversibles et la thermodynamique rationnelle auraient « beaucoup plus en commun que les deux camps veulent bien l'admettre » (Müller 2010 : 82).

La *thermodynamique prolongée* (« *extended thermodynamics* »), qui s'est développée à la fin du XX^e siècle, est influencée par la théorie cinétique des gaz et emprunterait les « meilleurs éléments de la

thermostatique, de la thermodynamique des processus irréversibles et de la thermodynamique rationnelle » (Müller 2010 : 95). À l'instar de la thermodynamique des processus irréversibles, l'objectif de la thermodynamique prolongée est la détermination d'un certain nombre de champs incorporés dans un vecteur de champs. La détermination de ces champs nécessite des équations empruntées à la mécanique et à la thermodynamique. Ces équations déterminent à leur tour des « relations constitutives » respectant généralement le critère de localité spatio-temporelle, en cela que les vecteurs dépendent uniquement des valeurs des champs pour un événement donné. Chaque solution des équations de champ constitue un processus thermodynamique. Mais il n'y a aucun matériau pour lequel les relations constitutives soient connues (Müller 2010 : 96). Le défi consiste donc à déterminer ces relations et trois outils sont disponibles, soit l'inégalité entropique, le critère de convexité et le principe de relativité. L'hypothèse de base est que ces champs déterminent des flux dits dissipatifs (flux de chaleur, tenseur de pression visqueuse, courant électrique) qui sont eux-mêmes déterminés par un ensemble de variables indépendantes de l'entropie. Certaines hypothèses quant aux caractéristiques de l'entropie sont alors nécessaires. Le principe combinant l'inégalité entropique ainsi que le critère de convexité permet d'établir un système d'équations de champ de forme hyperbolique symétrique, ce qui permet d'obtenir des vitesses finies de transmission de flux, comme des ondes de chocs. En effet, l'inégalité permet de restreindre la classe générique de champs devant être calculés afin de déterminer l'évolution temporelle des variables caractérisant les systèmes à l'étude. Ce projet est toujours en devenir (voir aussi Jou, Casas-Vàquez & Lebon 1988).

3.3 Mécanique statistique

La physique du XIX^e siècle a connu deux théories phares, soit la mécanique, issue des travaux de Isaac Newton et généralisée en formalisme par Pierre Simon de Laplace et Joseph-Louis Lagrange, et la thermodynamique, issue des travaux, comme on l'a vu plus tôt, de Sadi Carnot et développée par Rudolf Clausius, James Prescott Joule et William Thomson. Complémentaires à bien des égards, elles s'opposaient aussi parfois. L'opposition la plus fameuse, bien qu'elle ne fut pas la seule, est sans doute celle entre une théorie n'offrant pas de direction privilégiée au temps, en ce sens où les lois qu'elles présentent sont invariantes par renversement du temps, soit la mécanique, et une théorie dont l'un des principes fondamentaux stipule que les états physiques sont caractérisés par une quantité dont la valeur est supérieure à celle des états physiques antérieurs, la thermodynamique (Chapitre 4). La mécanique statistique est née de cette « tension » (Barberousse 2002 : 15).

Selon Brush (1977, in 2003 : 532), les trois aspects philosophiques de l'histoire de la mécanique statistique sont la réduction, la loi statistique et l'irréversibilité. En ce qui concerne le premier aspect, pour qu'il puisse y avoir réduction théorique, il faut qu'il y ait un niveau d'organisation dit inférieur suffisamment stable pour offrir une explication satisfaisante des phénomènes du niveau supérieur (Chapitre 2). C'est pourquoi il a d'abord fallu accepter, dans ce cas-ci, la thèse de l'existence de particules immuables et indestructibles, l'atomisme. Cette thèse, qui remonte bien entendu à l'Antiquité, avec entre autres Leucippe et Démocrite, a subi plusieurs modifications au cours des siècles et a fait l'objet de plusieurs spéculations avant de pouvoir être testée expérimentalement. On ne croit d'ailleurs plus, malgré l'importance de cette thèse, que les atomes sont insécables, indécomposables et indestructibles³⁹. Les travaux de Lavoisier et de John Dalton (1766-1844) se sont appuyés sur cette thèse afin de rendre compte de certaines réactions chimiques par les proportions des corps et ont ainsi ouvert la voie de la stœchiométrie, qui régit ces proportions. Mais plusieurs physiciens du XIX^e siècle d'obédience positiviste se sont opposés à la thèse atomiste sous prétexte que la science devait se limiter aux faits observables. Et des physiciens comme Boltzmann, l'endossant timidement, ont dû âprement s'en défendre⁴⁰. Il a fallu attendre le début du XX^e siècle avec les travaux d'Einstein sur le mouvement brownien, expliquant l'agitation de grains visibles par les mouvements individuels incessants des molécules invisibles, et d'une corroboration expérimentale en 1906 par Jean Baptiste Perrin (1870-

³⁹ « *If, in some cataclysm, all scientific knowledge were to be destroyed, and only one sentence passed on to the next generations of creatures, what system would contain the most information in the fewest words? I believe it is the atomic hypothesis (or atomic fact, or whatever you wish to call it) that all things are made of atoms.* » (Feynman 1996, in Ben-Naim 2008 : 316)

⁴⁰ Le titre de la biographie de Boltzmann par Cercignani (1998), *The man who trusted atoms*, est à ce propos évocateur.

1942) déterminant le nombre d'Avogadro, pour que la ligne tracée par le positivisme puisse être franchie. Sans entraîner un retentissant consensus cependant, le mouvement brownien fut accepté comme une manifestation directement observable des mouvements moléculaires et Perrin a pu affirmer quelques années plus tard la « réalité objective des molécules ».

Il a fallu ensuite, et c'est le deuxième aspect, apprécier les arguments probabilistes appliqués aux phénomènes naturels. En effet, à la Renaissance, la probabilité était associée à l'opinion en contraposition à la connaissance, vraie, absolue et certaine, et une opinion était « probable » si elle était « approuvée par une autorité » (Hacking 1975 : 28 ; Chapitre 5). Mais au XIX^e siècle, les probabilités commençaient leur incursion au sein d'explication scientifique. Par exemple, la méthode des moindres carrés, indépendamment élaborée par Legendre en 1805 et Gauss en 1809, permettait de comparer des données expérimentales, généralement entachées d'erreurs de mesure, à un modèle mathématique censé décrire ces données ; le darwinisme en 1859 fournissait une explication populationnelle où les individus ont une plus ou moins grande probabilité de survie dans un environnement donné en fonction d'une certaine propriété, la « *fitness* ». Et au XX^e siècle, les lois statistiques et probabilistes, de « pis-aller destinées à combler les insuffisances de nos moyens de connaissance, sont devenues des lois physiques autonomes » (Barberousse 2002 : 80). Cependant, au milieu du XIX^e siècle, le champ disciplinaire des probabilités était plutôt dans un état de changement et de confusion. Il est donc naturel que les travaux de Maxwell et Boltzmann en mécanique statistique aient reflété cet état. En effet, la méthode qu'a utilisée Maxwell en 1860 pour déterminer une fonction de distribution des vitesses des molécules dans un gaz, serait très semblable à celle de Quételet que Herschel a reprise en 1850 (Brush 1976 : 342). De plus, la position de Maxwell quant à la signification et à l'interprétation des probabilités est plutôt confuse, se situant entre les points de vue classique et fréquentiste. Tandis que Boltzmann a présenté une position équivoque par rapport à la signification des probabilités, oscillant entre celle d'une fréquence du temps et celle du nombre, et en se référant parfois à un ensemble de systèmes, parfois au volume en phase⁴¹.

Dans ce que les Ehrenfest (1912) ont appelé « la vieille formulation des investigations statistico-mécaniques », les états moléculaires, en particulier leur vitesse, sont considérées comme des variables stochastiques et les probabilités sont attribuées à ces états de mouvement. Ces probabilités sont elles-mêmes conçues comme des propriétés mécaniques de l'état global du système qu'est le gaz : elles sont censées représenter soit le *nombre* relatif de molécules possédant un état particulier, soit le *temps* relatif durant lequel une molécule a cet état (Uffink 2004 et 2007). Un mérite de cette approche est qu'elle

⁴¹ Uffink (2009 : 14-5) présente cinq interprétations différentes que Boltzmann aurait présentées dans ses écrits.

permet de prendre en compte les interactions entre molécules et qu'elle ne se restreint pas aux gaz mais peut aussi s'appliquer aux liquides et aux solides. Par contre, les probabilités deviennent plus « abstraites » car en étant attribuées aux états mécaniques du système total, elles ne sont plus déterminées par ces états mécaniques. Plutôt, les probabilités sont déterminées par un « ensemble », c'est-à-dire une collection fictive de répliques du système en question.

Enfin vient la question, soit le troisième aspect, de l'irréversibilité (Chapitres 4 et 5). La « tension » évoquée plus tôt est celle entre la réversibilité des lois de la mécanique classique et l'irréversibilité du second principe de la thermodynamique dont tente de rendre compte la mécanique statistique. Elle n'est donc pas étrangère au second aspect philosophique puisque cette difficile conciliation a trouvé un terrain d'entente, quoique discutable, avec la superposition d'une grille de lecture probabiliste appliquée à la pure description mécaniste.

3.3.1 James Clerk Maxwell

James Clerk Maxwell (1831-1879) est souvent considéré comme *le* physicien du XIX^e siècle (Einstein, in Pais 1982 ; Everitt 1983 ; Balibar 1999 ; Mahon 2003). Il est surtout reconnu pour avoir unifié l'électricité, le magnétisme et l'optique au sein d'un ensemble d'équations mathématiques, les « équations de Maxwell », créant ainsi la notion de champ électromagnétique. Il a su continuer la « tradition des savants des XVII^e et XVIII^e siècles, qui étaient tout autant physiciens, mathématiciens et philosophes », en sorte que ses recherches s'inscrivaient dans le « contexte d'une activité scientifique multiforme, et également dans un ensemble de réflexions situées à la frontière de la physique et de la métaphysique, en particulier au sujet de la réalité des atomes » (Barberousse (2002 : 70s). Il a beaucoup appris par lui-même et a fait montre d'une exceptionnelle dextérité dans le domaine des probabilités, un domaine dont on ne pouvait parler au milieu du XIX^e siècle de véritable « théorie ». Malgré des contributions remarquables à la théorie cinétique des gaz, aucun modèle de la théorie n'était, selon lui, en mesure de vraiment capturer la réalité des propriétés observées des gaz.

Son article « Illustrations de la théorie dynamique des gaz »⁴² (1860), sans doute inspiré d'un compte-rendu d'un livre de Quételet par Herschel (Brush 1976 : 184), s'inscrit autant dans la continuité que dans la nouveauté, car il marque la « renaissance de la théorie cinétique » (Uffink 2007 : 944), mais aussi « la première version d'une théorie suffisamment riche pour mériter le nom de 'mécanique

⁴² « *Illustrations of the Dynamical Theory of Gases* ». Texte intégral dans (Brush 2003) et traduction française partielle dans (Barberousse 2002).

statistique' » (Barberousse 2002 : 73). Maxwell y présente un *modèle* où un gaz est représenté comme un ensemble indéfini de petites sphères dures et parfaitement élastiques n'agissant l'une sur l'autre que lors des chocs qui induisent d'incessants changements dans les vitesses des « particules » du gaz modélisé. Son hypothèse fondamentale est celle à l'effet que les vitesses de ces particules ne tendent pas à s'égaliser, ce que peut laisser entendre le simple calcul des vitesses moyennes comme chez Clausius, mais plutôt à présenter une distribution statistique de valeurs données. Il calcule donc le « nombre moyen des particules dont les vitesses se trouvent entre des limites données après un grand nombre de collisions parmi un grand nombre de particules égales » (Maxwell 1860, in Barberousse 2002 : 87 ; aussi in Brush 2003 : 153), lequel est dénoté par

$$\mathcal{N}f(\mathbf{v}) d^3\mathbf{v},$$

où \mathcal{N} est le nombre total de particules et $f(\mathbf{v})$ est une fonction de distribution des vitesses, une fonction indéterminée de la vitesse. Si ce modèle permettait de déduire les propriétés macroscopiques des gaz, alors, selon Maxwell, on aurait établi une analogie physique importante, un « bon guide vers une explication authentique » (Barberousse (2002 : 73), car la simple cohérence du modèle est insuffisante à la validité de l'analogie.

Ce modèle repose sur deux importantes hypothèses : (i) l'*équiprobabilité* des directions possibles après collision, c'est-à-dire que la distribution $f(\mathbf{v})$ est sphériquement symétrique et ne peut dépendre que de la vitesse v ; (ii) l'*indépendance* des trois composantes de vitesse, c'est-à-dire qu'il est possible de factoriser $f(\mathbf{v})$ en fonction des trois composantes de vitesse orthogonales. Dans les mots de Maxwell (1860, in Brush 2003 : 153), les trois composantes de vitesse sont toutes à angle droit, donc orthogonales, et sont, par voie de conséquence selon lui, « indépendantes ». Cela signifie, par exemple, que les composantes de vitesse y et z sont statistiquement indépendantes de la composante x , et qu'en conséquence elles peuvent être traitées comme des variables stochastiques indépendantes. La satisfaction de ces hypothèses permet la résolution d'équations fonctionnelles menant à la fonction de distribution de type gaussien, celui-là même de la méthode des moindres carrés. Cet article de Maxwell serait le premier à caractériser l'état d'un gaz par une fonction de distribution f , et à qualifier $f(\mathbf{v}) d^3\mathbf{v}$ de *probabilité* (Uffink 2007 : 945). L'interprétation que propose Maxwell de cette probabilité est cependant plutôt confuse, se situant entre les points de vue classique et fréquentiste (Uffink 2007 ; Jodoin 2013a ; voir aussi le Chapitre 5)⁴³. Il obtient néanmoins des résultats importants, comme celui, plutôt « surprenant », montrant que la viscosité des gaz était indépendante de leur pression.

⁴³ Par exemple, Maxwell semble dériver la distribution f de considérations aprioriques, d'un desideratum de symétrie dérivée en quelque sorte du *principe de raison insuffisante* : s'il n'y a pas de raison de croire à une dépendance (entre certaines variables), alors il y a lieu de présumer une indépendance. Plus tard, en 1867 (in Brush 2003 : 219), il affirme : « [...] the assumption that the

Dans son article « Sur la théorie dynamique des gaz »⁴⁴ (1867) Maxwell prend ses distances des hypothèses de l'équiprobabilité et de l'indépendance lui ayant permis de calculer une fonction de distribution, ces hypothèses étant présentées comme « précaires ». Il traite dans cet article de ce que les Ehrenfest (1912) ont appelé la *Stoßzahlansatz*, un terme qui a fait école et qui signifie « hypothèse sur les collisions » : le nombre de collisions durant un temps dt , tel que $\mathcal{N}(\mathbf{v}_1, \mathbf{v}_2)$, dans lequel les vitesses initiales $\mathbf{v}_1, \mathbf{v}_2$ dans un élément $d^3\mathbf{v}_1 d^3\mathbf{v}_2$ sont changées en des vitesses finales $\mathbf{v}_1', \mathbf{v}_2'$ dans un élément $d^3\mathbf{v}_1' d^3\mathbf{v}_2'$ à l'intérieur d'un élément de volume spatial dV est proportionnel au produit du nombre de particules avec un vitesse \mathbf{v}_1 et du nombre de particules avec un vitesse \mathbf{v}_2 . Cette hypothèse, nous le verrons, joue un rôle primordial pour la validité du théorème H de Boltzmann, car elle en serait l'ingrédient, pour ainsi dire, d'irréversibilité (voir Chapitre 5).

Cependant, Maxwell n'a jamais tenté d'offrir une interprétation mécaniste du second principe de la thermodynamique, puisqu'il n'y croyait pas. Il serait plutôt le premier à proposer que le second principe de la thermodynamique soit de « nature statistique » (Pais 1982 : 61). Selon lui, le second principe avait « le même degré de vérité que l'énoncé à l'effet que si l'on jette un verre d'eau dans l'océan on ne peut récupérer le même verre d'eau »⁴⁵ (Maxwell 1867, in Klein 1970 : 86 ; 1973a : 62). Pour illustrer cette thèse, Maxwell a imaginé, dans une lettre à Peter Guthrie Tait en 1867 (Klein 1970), une expérience de pensée désormais célèbre où un être avec des capacités surhumaines, un « démon » selon l'expression consacrée de Thomson (1874), que n'endossait pas Maxwell, serait en mesure de discerner les molécules se déplaçant à grande vitesse dans un gaz de celles plus lentes. Cette expérience de pensée a été initiée par une discussion, parue dans *Nature* en janvier 1878 (in 1890b : 660s) dans son compte-rendu de l'ouvrage de Tait, un traité de thermodynamique, dont la « pierre de touche » doit être le second principe selon Maxwell. Cette discussion a trait à l'impossibilité d'offrir une interprétation mécaniste valable de ce principe, dont la vérité ne serait pas du même « ordre » que celle du premier principe (Maxwell 1878, in 1890b : 669). L'acuité avec laquelle Maxwell a su présenter ce problème est impressionnante. Selon lui, la différence, du point de vue moléculaire, entre les deux types de transfert d'énergie que sont la chaleur et le travail, est que les « mouvements et déplacements » associés au premier type échapperaient à nos méthodes d'observation, contrairement au second. De sorte que s'il était possible de suivre le mouvement des molécules, de la même façon que nous suivons le

probability of a molecule having a velocity resolved parallel to x lying between given limits is not in any way affected by the knowledge that the molecule has a given velocity resolved parallel to y . » Mais, avance Ellis (1850, in Uffink 2007 : 948), répondant à Herschel sur la même question, la simple ignorance ne fournit pas d'appui à quelque inférence que ce soit ; il affirme : « *It cannot be that because we are ignorant of the matter, we know something about it* ».

⁴⁴ *On the Dynamical Theory of Gases*. Texte intégral in (Brush 2003).

⁴⁵ « *The second law of thermodynamics has the same degree of truth as the statement that if you throw a tumblerful of water into the sea, you cannot get the same tumblerful of water out again.* »

déplacement d'un corps macroscopique comme un piston par exemple, alors la distinction entre la chaleur et le travail s'évanouirait. En conséquence, nous dit Maxwell (1878, in 1890b : 670), la vérité du second principe serait « statistique » et non pas « mathématique ».

Son raisonnement est le suivant. Si l'on considère le mouvement d'une seule molécule, on peut constater que son mouvement est irrégulier, qu'il change continuellement à chacune de ses collisions. En revanche, les propriétés dynamiques *moyennes* d'un petit groupe de molécules, bien que variant moins que celles d'une molécule individuelle, peuvent s'écarter considérablement de la valeur moyenne théorique puisque les membres individuels de ce groupe ne se comportent pas selon les lois décrivant le comportement moyen de l'ensemble. Le « second principe de la thermodynamique » est donc continuellement violé à l'échelle d'un groupe de molécules considérablement plus petit qu'un ensemble de molécules formant les corps macroscopiques observables. Plus ce groupe est grand, plus les déviations de la moyenne du groupe deviennent petites et moins fréquentes. Lorsque ce groupe est du même ordre de grandeur que celui d'un corps macroscopique observable, de l'ordre de 10^{23} molécules, « la probabilité d'un écart mesurable de la moyenne se produisant à l'intérieur d'un nombre fini d'années devient si petit qu'il peut être considéré en pratique comme une impossibilité »⁴⁶ (Maxwell 1878, in 1890b : 670). Bien que peu probable, cette impossibilité ne peut faire l'objet d'une « certitude absolue », et elle ne peut être, par conséquent, une impossibilité stricte. Il est donc vain, selon lui, de tenter de déduire le « second principe de la thermodynamique » de principes « purement dynamiques ».

3.3.2 Ludwig Boltzmann

Ludwig Boltzmann (1844-1906) a été le premier à tenter d'offrir une preuve mécanique du second principe de la thermodynamique. La motivation pour cette entreprise fastidieuse découle sans doute de sa préférence pour des systèmes axiomatiques généralisables, hypothético-déductifs, sans adopter un réalisme naïf quant à l'existence des atomes dans climat philosophique dominé par l'empirisme phénoménologique et l'inductivisme (Cercignani 1998 ; Reiter 2007). En fait, Boltzmann s'est intéressé

⁴⁶ « [...] *the probability of a measurable variation from the mean occurring in a finite number of years becomes so small that it may be regarded as practically an impossibility* ».

de près à la philosophie, sans toutefois apprécier pleinement cette discipline, mais peut à bien des égards être considéré comme un philosophe des sciences à part entière⁴⁷.

La première tentative de Boltzmann de trouver un équivalent mécanique du principe entropique se trouve dans un article de 1866, « Sur la signification mécanique de la deuxième loi de la thermodynamique ». À cette époque il ne connaissait pas les articles de Maxwell de 1860 et 1867 (Uffink 2007 : 952). Cependant, Rankine a publié un an plus tôt en 1865, et seulement quelques mois après l'article de Clausius introduisant le concept d'entropie, un article intitulé « Du second principe de la thermodynamique »⁴⁸, dont l'objectif explicite était de « donner une preuve élémentaire de l'hypothèse selon laquelle le second principe de la thermodynamique découle de la supposition que la chaleur sensible consiste en une sorte de mouvement moléculaire régulier dans un espace clos »⁴⁹. Mais cet article est plutôt incomplet et préliminaire, où l'auteur esquisse, sans mathématique, les propriétés d'une « fonction thermodynamique », qui serait équivalente à l'entropie.

La stratégie de Boltzmann dans son article « Études sur l'équilibre de la force vive entre des points matériels en mouvement »⁵⁰ (1868a), consistait en une exploration systématique de tous les aspects mathématiques du mouvement de grands systèmes dans le but de donner des *fondements analytiques* à la théorie mécanique de la chaleur. Il y reconnaît les avancées de Maxwell mais affirme que son argumentation est « difficilement compréhensible en raison de sa grande concision, et encore plus difficilement compréhensible à cause d'une erreur d'impression [...] » (in Barberousse 2002 : 151). Contrairement à Maxwell, Boltzmann adopte une méthode géométrique mais aussi une interprétation sensiblement différente des probabilités :

Soit $\phi(c)dc$ la somme des moments du temps, prise pendant un temps très long, pendant lesquels la vitesse d'un cercle quelconque est comprise entre c et $c + dc$, divisée par ce temps très long, c'est-à-dire la probabilité que c soit comprise entre c et $c + dc$, et soit N le nombre moyen des cercles dont les centres se trouvent dans une unité de surface, alors

$$N\phi(c)dc$$

⁴⁷ Boltzmann a donné plusieurs conférences sur des sujets proprement philosophiques, dont l'une portait d'abord le titre « *Proof that Schopenhauer is a stupid, ignorant philosophaster, scribbling nonsense and dispensing hollow verbiage that fundamentally and forever rots people's brains* » (Cercignani 1998 ; Müller 2007), avant d'adopter un titre plus « conventionnel ».

⁴⁸ « *On the second law of thermodynamics* ». Cercignani (2001 : 29) cite un autre article de Rankine (1867), « Sur la nécessité de vulgariser la seconde loi de la thermodynamique », comme étant une tentative de trouver un équivalent mécanique du principe entropique. Mais je ne vois rien dans cet article qui puisse se rapprocher d'une telle tentative, car comme son titre l'annonce, l'article est plutôt un plaidoyer en faveur d'une plus grande diffusion pédagogique de ce principe, négligé selon lui par rapport au premier principe.

⁴⁹ « *Über die Mechanische Bedeutung des Zweiten Hauptsatzes der Wärmetheorie* ».

⁵⁰ « *Studien über das Gleichgewicht des lebendigen Kraft zwischen bewegten materiellen Punkten* ». Traduction française partielle dans (Barberousse 2002).

est le nombre de cercles par unité de surface dont les vitesses sont comprises entre c et $c + dc$. (1868, in Barberousse 2002 : 151).

Ainsi, $\phi(c)dc$ est introduit comme le *temps* relatif durant lequel un cercle donné possède une certaine vitesse ; mais, du même souffle, cette quantité est identifiée comme le *nombre* relatif de cercles ayant cette vitesse. Boltzmann appliquerait donc deux significations à la même fonction et il a maintenu cette confusion entre ces deux interprétations tout au long de ses écrits (Uffink 2007 : 953 ; Jodoin 2013a). De plus, il s'appuie, à l'instar de Maxwell avant lui, sur la *Stoßzahlansatz* et il annonce un théorème général, ce qui deviendra l'*hypothèse ergodique*⁵¹. Cette hypothèse, placée à la toute fin de son article, vient indiquer que la distribution des vitesses des molécules est nécessairement de type maxwellien sauf dans le cas où les variables dynamiques « reviennent à [leurs] valeurs initiales après une période déterminée, sans avoir pris au cours de ce mouvement toutes les valeurs compatibles avec le principe de la force vive [équation de l'énergie] » (Boltzmann 1868, in Barberousse 2002 : 165). Elle allait susciter de grands débats, qui se poursuivent encore, quant aux fondements de la mécanique statistique (voir Brush 1976 : 241-2 ; Uffink 2007 : 955-6 ; Jodoin 2011a ; Chapitre 5).

Dans son article « Nouvelles études sur l'équilibre de la chaleur parmi des molécules gazeuses »⁵² (1872), Boltzmann obtient deux résultats d'importance, désormais connus comme l'*équation de Boltzmann* et le *théorème H*. Il souhaite prouver que, peu importe la distribution initiale (f) des états, elle doit inévitablement approcher la distribution développée par Maxwell, et ensuite s'y maintenir, ce qui correspond à un analogue mécanique statistique de l'équilibre. Son raisonnement quant à sa méthodologie est le suivant : selon la théorie mécanique de la chaleur, les molécules d'un gaz sont en mouvement constant et rapide ; et en raison de leur nombre prodigieux, nous ne pouvons observer ni calculer rien d'autre que la valeur moyenne des propriétés de l'ensemble ; or la détermination des valeurs moyennes est l'objet du calcul des probabilités ; donc, les problèmes de la théorie mécanique de la chaleur sont aussi des problèmes du calcul des probabilités. Cependant, Boltzmann précise que

⁵¹ Les questions de l'origine et de l'étymologie du terme « *ergode* » ont donné lieu à des débats toujours actifs aujourd'hui. Brush (1976 : 364) soutient que Boltzmann a utilisé ce terme pour désigner un ensemble stationnaire caractérisé par la distribution microcanonique. De ce point de vue, il semble qu'il n'ait rien à voir avec l'hypothèse ergodique, comme l'ont soutenu les Ehrenfest (1912 : 89). Ces derniers, en effet, affirment que Boltzmann l'aurait introduit dans un article de 1887 à partir des mots grec « *εργον* » (énergie, projet) et « *οδος* » (chemin, rue). Gallavotti (1994) soutient que le terme viendrait plutôt des mots « *εργον* » et « *ειδος* » en référence à une « vraie nature », au sens platonicien du terme. Mais son argumentation ne convainc par Uffink (2007) et ne me convainc pas non plus. En fait, il semble que ce soit Clausius (1870) qui soit le premier à se référer au mot « *εργον* » pour définir l'*ergal*, soit la valeur moyenne de la *vis viva* dans son *théorème du viriel*. La propension de Clausius à former des néologismes à partir de mots grecs est bien connue et il est étonnant que ces auteurs – les Ehrenfest, Brush, Gallavotti et Uffink – n'en aient pas fait mention. Sans doute parce que cet article de Clausius reprend en quelques sortes des résultats déjà obtenus par Boltzmann.

⁵² « *Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen* ». Traduction française partielle dans (Barberousse 2002) ; traduction anglaise dans (Brush 2003).

l'utilisation de probabilités n'entache pas d'incertitude les résultats du raisonnement, lesquels découlent rigoureusement des prémisses établies.

Il construit une équation différentio-intégrale d'évolution de f_i , en prenant la différence du nombre de collisions durant un intervalle de temps dt et du nombre relié aux collisions pour ainsi dire inverses déjà calculés par Maxwell. Il doit alors supposer, d'une part, que la force entre deux particules sphériques est fonction de leur distance mutuelle, d'autre part que cette force agit dans le sens des lignes de centre, et enfin que l'action et la réaction de cette force sont égales. Le résultat est l'*équation de Boltzmann*, une équation qui décrit l'évolution temporelle de f selon des conditions initiales données (voir Brush 1976 : 600 ; Cercignani 1998 : 89 ; Uffink 2007 : 964 ; Müller 2007 : 95). Avec cette équation Boltzmann est en mesure de prouver non seulement la *stationnarité* de la distribution de Maxwell mais aussi son *unicité*, c'est-à-dire que cette distribution serait la seule à satisfaire le desideratum d'une invariance temporelle. Ce résultat suggère évidemment que la stabilité la distribution de Maxwell, son invariance temporelle, représentée par un taux de variation nul, équivaut à une situation d'équilibre. L'étape suivante dans la démonstration de Boltzmann consiste à prouver qu'il y ait une *tendance* vers une distribution maxwellienne. L'avantage de l'équation de Boltzmann est donc de présenter l'évolution de f_i (soit la fonction de distribution de probabilité pour une particule donnée en un certain point de l'espace des phases réduit) lorsque les conditions initiales de f_i sont données.

Supposant alors que l'équation qui porte son nom et que la *Stoßzahlansatz* soient valides pour tous les temps, Boltzmann est en mesure de prouver que la quantité (en notation moderne)

$$H[f_i] = \int f_i(\mathbf{v}, t) \ln f_i(\mathbf{v}, t) d^3\mathbf{v},$$

décroît de manière monotone dans le temps, de sorte que

$$dH[f_i]/dt \leq 0 ;$$

et cette inégalité devient une égalité si et seulement si la distribution des vitesses est de type maxwellien, c'est-à-dire

$$dH[f_i]/dt = 0, \text{ si et seulement si } f_i(\mathbf{v}) = A \exp(-Bv^2),$$

ce qui constitue en substance le *théorème H* (Brush 1976 : 238 ; Zeh 1989 : 35 ; Uffink 2007 : 965 ; Brown, Myrvold & Uffink 2009 : 175). Autrement dit, toute distribution évoluerait vers une distribution de type maxwellien puisque $H[f_i]$ décroît toujours jusqu'à atteindre un minimum stationnaire dont la distribution est justement de ce type⁵³. Cette quantité $H[f_i]$, qui est plutôt représentée par E (pour

⁵³ La preuve de la décroissance de $H[f_i]$ est plutôt directe puisqu'elle repose sur le simple fait mathématique que la quantité $(a - b) \log(b/a)$ est toujours négative si a et b sont des réels positifs.

« entropie » selon les Ehrenfest 1912 : 84, n. 53) dans l'article original, serait le fruit de « conjectures informées » (Brush 1976 : 600), d'une inspiration issue d'essais et d'erreurs. Le point important ici est de remarquer que la fonction $H[f_i]$ présente un comportement *analogue* à celui de la fonction S de Clausius. Boltzmann (1872, in Brush 2003 : 263, italiques originaux) infère une « preuve » de cette identification d'un analogue : « [p]uisque $[H]$ est reliée de près à l'entropie thermodynamique à l'état final d'équilibre, notre résultat est équivalent à une preuve que l'entropie doit toujours *augmenter* ou demeurer constante, et il fournit ainsi une interprétation microscopique du second principe de la thermodynamique »⁵⁴.

L'apport théorique de Boltzmann avec cet article est considérable. Il est en effet remarquable, nonobstant certains problèmes théoriques et épistémologiques qui ont suscité et suscitent de vives discussions, qui sont souvent l'apanage des grands concepts, d'exprimer l'entropie S en termes d'une fonction de distribution, soit $S = -kH$. Il importe d'ailleurs de noter que l'équation $H[f_i]$ serait la première à décrire l'« évolution temporelle d'une probabilité » (Cercignani 1998 : 97), ce qui n'est pas sans rappeler, en mécanique quantique, l'équation de Schrödinger, établie cinquante ans plus tard. Ce qu'apporte Boltzmann avec ce théorème est double, puisqu'il fournit non seulement une *généralisation* du concept d'entropie aux états de non-équilibre, représentés par des distributions qui ne sont *pas* du type maxwellien, mais aussi l'*affirmation* à l'effet que l'entropie de non-équilibre définie avec la quantité $H[f_i]$ s'accroît de manière monotone, alors que la fonction de distribution décrivant l'état d'un gaz isolé évolue d'un état de non-équilibre à un état d'équilibre. Toutefois, autant la question de la valeur de cet analogue, car Boltzmann ne parle pas explicitement d'échange de chaleur comme le fait Clausius, que celle de sa portée, c'est-à-dire comme véritable preuve que l'entropie augmente, demeurent discutables. Mais l'enjeu le plus important concerne la caractéristique d'irréversibilité exprimée par le théorème H avec la décroissance de la quantité $H[f_i]$.

Un problème évident dont fait l'objet ce théorème concerne l'incompatibilité, apparente ou intrinsèque, entre, d'une part, cette caractéristique d'irréversibilité et, d'autre part, la mécanique newtonienne, à partir de laquelle ce théorème devrait être dérivée en toute rigueur logique, dont les lois sont absolument réversibles, c'est-à-dire qu'elles sont tout aussi valides, en termes simples, lors d'un remplacement de t par $-t$. Ce problème, qui en implique d'autres et qui prend des allures de véritable *paradoxe*, est rien de moins que gigantesque, en ceci qu'il fait intervenir toute une pléthore d'enjeux physiques, épistémologiques et métaphysiques, dont ceux, et non des moindres, d'une réduction théorique et de la direction du temps (Chapitres 4 et 5). Il n'est donc pas étonnant qu'il ait suscité

⁵⁴ « Since E is closely related to the thermodynamic entropy in the final equilibrium state, our result is equivalent to a proof that the entropy must always increase or remain constant, and thus provides a microscopic interpretation of the second law of thermodynamics. »

autant de débats. L'un des premiers à le soulever est sans doute Thomson (1874) dans un article, qui n'attira pas l'attention, où il affirme que les phénomènes irréversibles de la « dissipation de l'énergie » pourraient être « empêchés » par un arrangement adéquat de « démons de Maxwell ». Mais les deux objections les plus célèbres sont celles de Josef Loschmidt (1821-1884), ami de longue date et mentor de Boltzmann, qu'on appelle l'*objection de la réversibilité*, et celle de Ernst Zermelo (1871-1953), soit l'*objection de la récurrence*. Voici brièvement de quoi il retourne (voir Chapitre 5).

L'expression « paradoxe de la réversibilité » est due aux Ehrenfest (1907) et l'objection qui en est au centre provient de plusieurs discussions entre Maxwell, Tait et Thomson, ayant introduit d'ailleurs la notion, déjà discutée, de démon de Maxwell (Brush 1976 : 602). Loschmidt (1876) y fait allusion alors qu'il tente, de manière plutôt absconse, de déconstruire le « terrible nimbus » du second principe semblant impliqué la « mort thermique de l'univers ». Ce que met en exergue l'énoncé de ce paradoxe est la conclusion contradictoire du caractère irréversible exprimé par le théorème H obtenue de prémisses stipulant la réversibilité des lois de la mécanique classique. Ainsi, pour chaque distribution de vitesses compatible avec des paramètres macroscopiques (thermodynamiques) donnés, il existe une distribution aussi compatible avec ces paramètres mais où les vitesses sont inversées. Cela signifie que pour toute trajectoire dynamique compatible avec l'évolution d'un système thermodynamique d'un état initial à un état final, et présentant des caractéristiques apparemment irréversibles comme dans le cas de la fonte d'un glaçon, il existe une trajectoire dynamique inverse compatible avec l'évolution de ce système thermodynamique de son état final vers son état initial.

Dans son article « Sur la relation entre un théorème mécanique général et la seconde loi de la théorie mécanique de la chaleur »⁵⁵ (1877a), Boltzmann apporte une réponse à l'objection de son ami Loschmidt et reconnaît explicitement que toute tentative de prouver l'inégalité de Clausius ($\int \delta Q/T \leq 0$) à partir des seules lois de la dynamique et sans égards aux conditions initiales doit être « futile ». Son argumentation repose sur deux éléments qui annoncent un changement conceptuel chez Boltzmann : la nature *probabiliste* du second principe de la thermodynamique et l'importance des *conditions initiales*. Ainsi, il considère la permutation des molécules individuelles comme impliquant un nouvel arrangement, ce qu'il appellera plus tard une « complexion » (« *Komplexion* »), souvent appelée aujourd'hui *micro-état*. Le nombre relatif de complexions compatibles, circonscrit par des paramètres macroscopiques donnés, indiquerait sa probabilité. En sorte que Boltzmann ne définit plus les probabilités à partir de la distribution des vitesses mais plutôt il leur attribue une probabilité d'après une hypothèse d'équiprobabilité, à l'instar du principe laplacien de raison insuffisante. Or, il se trouve

⁵⁵ *Über die Beziehung eines Allgemeine mechanischen Satzes zum zweiten Hauptsätze der Warmtheorie*. Traduction anglaise dans Brush (2003).

qu'il existe beaucoup plus de distributions uniformes, représentant des états d'équilibre, que de distributions non uniformes, représentant des états de non-équilibre. Il est donc beaucoup plus probable, affirme Boltzmann, qu'un système se trouve dans un micro-état évoluant selon les lois dynamiques vers un état caractérisé par une distribution uniforme qu'autrement. C'est-à-dire que l'état initial dont parle Loschmidt, qui mènerait à un comportement allant à l'encontre de la thermodynamique, est très peu probable. En d'autres mots, selon cette ligne de pensée, si l'on choisit un état initial au hasard, l'on est à peu près certain que son entropie va augmenter. On constate donc que Boltzmann s'appuie fortement sur la théorie des probabilités. Mais cette dernière est aussi neutre quant à la direction des processus que peut l'être la mécanique. L'hypothèse du *Stoßzahlansatz*, comme l'ont souligné des articles animant un débat dans les pages de *Nature* en 1894, serait à l'origine de cette irréversibilité. (Pour une discussion plus détaillée, voir Chapitres 4 et 5.)

Quelques mois plus tard, dans son article « Sur la relation entre un théorème mécanique général et la seconde loi de la théorie mécanique de la chaleur »⁵⁶ (1877b), Boltzmann poursuit avec ce changement conceptuel par rapport à ses positions antérieures, notamment à l'effet qu'une preuve du second principe de la thermodynamique est possible sur la *seule* base du calcul des probabilités. C'est dans cet article que Boltzmann présente le lien mathématique entre entropie et probabilité, ce qui deviendra plus tard l'*entropie combinatoire de Boltzmann* :

$$S_{B,c} = k \ln W,$$

une formule inscrite sur sa pierre tombale, bien qu'il ne l'ait jamais écrite explicitement, ce que l'on doit plutôt à Planck. Boltzmann, dans ce qu'il est convenu d'appeler l'*argument combinatoire*, procède à la définition de cette entropie de la façon suivante. Il procède d'abord à une partition en petites « cellules » de l'espace des états possibles des molécules composant un gaz, qu'on appelle espace des phases ou des configurations, où chacune des cellules possède un même « volume » dans cet espace mathématique et une énergie précise. Il définit ensuite une « distribution d'état » (« *Zustandsverteilung* »), soit le nombre de molécules pour chacune des cellules définies, en quelques sortes le nombre de « combinaisons » possibles. Si l'on admet qu'une permutation des molécules dans ces cellules donne une nouvelle distribution d'état, alors chacune de ces distributions d'état ainsi distinguées forme ce qui a été nommé une « complexion » (« *Komplexion* »). Pour Boltzmann, à l'instar du jeu de loto où chaque séquence de chiffre (par exemple, 12345) est aussi probable qu'une autre, chaque complexion est aussi probable, donc équiprobable. Or, divisant le nombre de complexions associées à une distribution d'état donnée par le nombre total de complexions on obtiendrait la *probabilité* (« *Wahrscheinlichkeit* »), W ,

⁵⁶ « Über die Beziehung eines Allgemeine zwischen dem zweiten Hauptsatze der mechanischen Warmtheorie und der Warschenlichkeitsrechnung respektive dem Sätzen über das Warschenlichgewicht ».

de cette distribution d'état. De plus, comme sa démonstration est statique, c'est-à-dire qu'elle ne concerne pas l'évolution des systèmes, il précise qu'un système de corps (« *System von Körpern* ») va toujours d'un état improbable à un état plus probable.

Zermelo (1896a) a soulevé une seconde objection, toute aussi puissante sinon plus que la précédente, à partir d'un argument faisant appel au théorème de la récurrence de Poincaré (1890). Ce théorème implique que, dans le cas de l'évolution d'un système à plusieurs degrés de liberté comme un gaz dont le comportement des composantes est régi par les lois de la mécanique classique et par le principe de la conservation de l'énergie, pour un flot hamiltonien d'un espace des phases borné, chaque état doit retourner exactement à son état initial et ce, avec une récurrence arbitraire, excepté pour un ensemble d'états singuliers. Par conséquent, toute fonction continue sur l'espace des phases ne peut s'accroître de manière monotone dans le temps, comme l'affirme pourtant le théorème *H*, car s'il y a un accroissement fini, alors il y a aussi une décroissance lorsque l'état initial se reproduit, d'où la récurrence. En d'autres mots, tout système mécanique conservatif dont l'espace des phases est fini est nécessairement quasi-périodique. Pour que l'état initial se reproduise, que la récurrence soit effective, il n'est donc « plus nécessaire d'avoir la vue fine, la présence d'esprit, l'intelligence et l'adresse du démon de Maxwell, il suffi[t] d'un peu de patience » (Poincaré 1893 : 536). En conséquence, il serait « impossible de prouver » que la distribution de type maxwellien soit obtenue comme « résultat stationnaire final » (Zermelo 1896, in Brush 2003 : 391).

La riposte de Boltzmann vient dans sa « Réplique à la remarque de Zermelo sur la théorie de la chaleur »⁵⁷ (1896a). En bref, il soutient que, malgré la validité du théorème de Poincaré, son application à la théorie cinétique est inexacte. En effet, selon lui, la critique de la contradiction avec l'expérience est inopérante puisque, sur une période de temps suffisamment longue, les états menant à un état d'équilibre se présentent dans une écrasante majorité. Ainsi, la récurrence demeure possible mais sur une période de temps extrêmement longue. Par exemple, le temps pour une récurrence de l'état initial d'un centimètre cube d'air est estimé à $10^{\mathcal{Z}}$ secondes, où $\mathcal{Z} = 10^{19}$, un temps largement supérieur à l'âge estimé de l'univers. L'affirmation de Poincaré apparaît donc comme un euphémisme des plus optimistes. En affirmant sur une base théorique que l'état du gaz doit se reproduire, Zermelo, affirme Boltzmann (1896a, in Brush 2003 : 397), se comporte « comme un joueur de dés qui, ayant calculé que la probabilité d'une séquence de mille 'un' n'est pas nulle, en conclut que son dé doit être pipé puisqu'il n'a pas observé une telle séquence ». Par conséquent, la récurrence est au-delà de toute expérience humaine et ne peut donc invalider l'approche mécanique qu'il propose.

⁵⁷ « *Entgegnung auf die wärmetheoretischen Betrachungen des Hrn. Zermelo.* » Traduction anglaise dans (Brush 2003)

3.3.3 Josiah Willard Gibbs

Le livre de Josiah Willard Gibbs (1839-1903) *Principes élémentaires de la mécanique statistique* (1902) marque la naissance de la mécanique statistique au sens strict, soit comme une théorie « cohérente et systématique » (Uffink 2007 : 992) offrant « fondation rationnelle » (*ibid.*) à la thermodynamique. Selon Barberousse (2002 : 198), à lire la préface de Gibbs, on constate « de façon particulièrement frappante, combien sont importantes pour la physique statistique les considérations épistémologiques, ou encore le regard réflexif et critique sur les conditions de production de la connaissance scientifique ». Seul Einstein semble avoir surpassé Gibbs dans ce « genre d'acuité théorique permettant d'obtenir des résultats probants de manière isolé » (Cropper 2001 : 106).

En caractérisant l'approche de Gibbs, plusieurs auteurs, dont Sklar (1993) et Uffink (2007), la qualifient de « prudente ». En fait, son approche repose sur des *analogies* entre des ensembles statistiques de molécules et certains résultats de la thermodynamique : étant donné certaines hypothèses quant à la distribution des variables canoniques, l'équation de cette distribution prend la *forme* d'une équation thermodynamique qui peut alors être considérée comme son analogue. Il s'est bien gardé d'offrir des justifications à ses associations ou un modèle microscopique de la matière. Aussi n'a-t-il jamais prétendu avoir *réduit* la thermodynamique à la mécanique statistique. Son approche est tout de même novatrice. Les seuls ingrédients du « schème logique » (Uffink 2007 : 992) de Gibbs sont la mécanique et les probabilités, lesquels sont à l'œuvre dans ce qu'on appelle la *méthodologie des ensembles*. Celle-ci permet de traiter certaines propriétés des atomes individuels, comme l'énergie, comme des variables indépendantes et de calculer des moyennes sur ces variables.

On peut alors imaginer un grand nombre de systèmes de même nature mais qui diffèrent selon les configurations et vitesses qu'ils ont à un instant donné, pas seulement de façon infinitésimale, mais de sorte à embrasser toutes les combinaisons concevables de configurations et de vitesses. Et ici on peut poser le problème, non pas de suivre la séquence des configurations d'un système particulier, mais de déterminer comme le nombre total de systèmes sera distribué parmi les diverses configurations et vitesses concevables à n'importe quel moment, lorsque cette distribution est déjà donnée pour un certain temps.⁵⁸

⁵⁸ « We may imagine a great number of systems of the same nature, but differing in the configurations and velocities which they have at a given instant, and differing not merely infinitesimally, but it may be so as to embrace every conceivable combination of configurations and velocities. And here we may set the problem, not to follow a particular system through its succession of configurations, but to determine how the whole number of systems will be distributed among the various conceivable configurations and velocities at any required time, when the distribution has been given for some one time. » (Gibbs 1902 : vii ; traduction libre).

Les systèmes d'un ensemble de Gibbs n'interagissent pas, il s'agit en quelque sorte d'un assemblage imaginaire. Gibbs définit quelques cas particuliers de ces ensembles, soit l'ensemble « microcanonique », décrivant des systèmes isolés de leur environnement et donc à énergie constante, l'ensemble « canonique », où le volume, la température et le nombre de molécules sont constants, et l'ensemble « grand-canonique » dont l'énergie et le nombre de molécules peuvent varier mais où le potentiel chimique demeure constant.

En définissant ainsi un ensemble de « sous-systèmes » possibles, soit ces « combinaisons concevables de configurations et de vitesses » en fonction de certains paramètres donnés, c'est-à-dire mesurés ou mesurables, et de leur distribution dans l'espace des possibles qu'est l'espace des phases, il va de soi que Gibbs allait définir la *probabilité* de ces « sous-systèmes ». Celle-ci est établie, dans son formalisme, par le produit d'une fonction de densité de probabilité, une fonction mathématique exprimant la distribution des probabilités des variables aléatoires, et du « volume » d'une région de l'espace des phases, ce qui donne le nombre relatif de systèmes dans l'ensemble en question dont le micro-état se trouve dans cette région. Bien que Gibbs soit resté plutôt « évasif » (Uffink 2007 : 993) quant aux diverses interprétations des probabilités, ce qui est naturel étant donné que leurs traitements systématiques sont pour la plupart subséquents à sa mort, cet appel au nombre relatif laisse entendre une interprétation fréquentiste (Chapitre 5). Ainsi, la probabilité qu'un système ait telle caractéristique particulière est déterminée par les paramètres définissant l'ensemble auquel il appartient ; par exemple, dans l'ensemble microcanonique, la probabilité qu'un système ait une certaine valeur d'énergie est proportionnelle à un facteur, le « facteur de Boltzmann », faisant intervenir l'énergie et la température. La densité de probabilité est une fonction qui dépend des variables mécaniques généralisées, la position et la vitesse, et son évolution est déterminée par les lois de la mécanique, plus précisément les équations hamiltoniennes du mouvement. Or, le comportement de cette densité représente en quelque sorte le comportement du système thermodynamique, celui susceptible d'être observé et mesuré. C'est pourquoi une densité constante dans le temps représente selon Gibbs la condition d'un *équilibre statistique*, un *analogue* de l'*équilibre thermique*.

Gibbs parle ainsi d'analogues en mécanique statistique de quantités thermodynamiques comme la température et l'entropie. Bien que, selon lui, une « définition mécanique » (Gibbs 1902 : 165) de ces deux quantités permettrait d'ailleurs d'établir une « fondation apriorique » (*ibid.*) des principes de la thermodynamique, il n'utilise pas un raisonnement déductif pour les obtenir. Gibbs présente au moins trois définitions différentes de l'entropie statistique. Il affirme ainsi avoir trouvé une équation différentielle dont la « forme est identique » (1902 : xi) à celle de l'« équation thermodynamique fondamentale » dans le cas de l'ensemble canonique :

$$d\langle H \rangle = \theta dS_G - \sum \langle A_i \rangle da_i,$$

où $\langle H \rangle$ est l'espérance mathématique de l'hamiltonien dans l'ensemble canonique, θ est le module de l'ensemble (et correspond à la température thermodynamique), S_G (représentée par η dans le (1902)) est ce qu'il appelle l'« index moyen de probabilité » et $\langle A_i \rangle$ représente les « forces généralisées (Gibbs 1902 : 44). Cette analogie lui permet alors d'identifier, par association de formes, l'« index moyen de probabilité » avec l'entropie thermodynamique. Une formule similaire peut être obtenue pour les ensembles microcanonique et grand-canonique.

Pour définir l'entropie S_G , Gibbs fait appel à ce qui s'apparente à un élément de volume d'une région de l'espace des phases (« *extension-in-phase* »), en ceci qu'il s'agit du produit des éléments infinitésimaux des variables mécaniques généralisées pour chacun des degrés de libertés, autant dire pour chacune des particules, ainsi que la densité en phase, une fonction de ces variables. Le produit de cet élément de volume et de cette densité donne le nombre de systèmes d'un ensemble se trouvant dans cette région de l'espace des phases. Et en divisant ce résultat par le nombre total de systèmes de l'ensemble, ce que l'on suppose possible lorsque la densité *et* ce nombre tendent vers l'infini, on obtient la *probabilité* qu'un système quelconque se trouve dans cette région, dans certaines limites de l'espace des phases. Gibbs (1902 : 16) affirme que la quantité, qu'il nomme l'« index de probabilité » et qui apparaît dans l'équation thermodynamique fondamentale présentée plus tôt, se définit par le logarithme naturel de cette probabilité⁵⁹. Gibbs (1902 : 128 et 176) présente deux autres définitions de l'entropie pour l'ensemble microcanonique, soit celle qui se définit par le logarithme naturel du volume d'une région de l'espace des phases dont l'énergie est inférieure ou égale à une certaine valeur d'énergie, qu'on appelle désormais l'« entropie de volume », et celle qui se définit de la même façon mais où l'énergie est strictement égale à une certaine valeur d'énergie, qu'on appelle désormais l'« entropie de surface ». En outre, l'argumentation, plutôt directe en ce qu'elle s'appuie pour l'essentiel sur la reconnaissance de ressemblance formelle, et menant à ces définitions, n'exclut pas la possibilité de trouver d'autres définitions tout aussi acceptables. C'est pourquoi « il pourrait y avoir d'autres quantités pouvant être considérées comme [...] l'entropie » (Gibbs 1902 : 169).

Il y a évidemment plusieurs questions et plusieurs problèmes qui demeurent avec ces définitions, certains ayant été déjà soulevés par Gibbs lui-même. En voici un aperçu. D'abord, la justification du choix des ensembles statistiques présentés par Gibbs pour décrire des situations d'équilibre est pour ainsi dire inexistante, si ce n'est la raison de leur simplicité. Des travaux plus anciens et d'autres subséquents, comme ceux de Boltzmann (1868) et d'Einstein (1902), identifient l'ensemble

⁵⁹ L'équation de cette définition n'est pas explicitée chez Gibbs, mais elle recevra une définition formelle au Chapitre 5.

microcanonique comme la *seule* distribution stationnaire pour des systèmes isolés *si* la condition d'ergodicité est satisfaite, une hypothèse qui n'est pas sans poser problèmes. Ensuite, l'insistance sur les analogies formelles tend à faire oublier la signification physique des différents concepts d'entropie, ce que doit établir une discussion sur les unités attribuer aux quantités physiques. Gibbs soutient que ses concepts d'entropie, définis par le logarithme naturel du volume en phase, sont affectés par les unités de masse et de temps, ce qui peut se corriger par l'application d'une constante adéquatement choisie, mais puisque seules les différences d'entropie ont une signification physique, le choix de cette constante semble « immatériel » (Gibbs 1902 : 206). Mais dans le cas du mélange de plus d'une quantité de la même substance, on risque d'y perdre la propriété, importante au demeurant, de l'extensivité, un problème, désigné comme le « paradoxe de Gibbs », qu'il tente de résoudre en distinguant l'« entropie générique » et l'« entropie spécifique ». Enfin, si l'état du système est caractérisé par la densité de probabilité, alors le changement dans le temps de cette densité est censé représenter l'évolution dans le temps du système. Et si cette densité change, alors, par la définition précédente, l'entropie du système va aussi changer. Mais il y a un théorème mathématique, le théorème de Liouville, qui stipule que lorsque l'évolution temporelle d'une densité est gouvernée par les équations hamiltoniennes et dans certaines conditions qui sont remplies dans le contexte de la mécanique classique, alors celle-ci demeure constante. Évidemment, si la densité des systèmes de la mécanique statistique demeure constante dans le temps, alors l'entropie de Gibbs est aussi constante, et par conséquent elle ne peut décrire l'évolution des systèmes vers l'équilibre. (Plus de détails au Chapitre 5.)

Gibbs a donc apporté des outils conceptuels très importants en physique, fondant ainsi une nouvelle discipline scientifique, la mécanique statistique. Son style « prudent », sa conscience des enjeux épistémologiques épineux qu'elle mettait en place malgré sa réticence à s'y engager pleinement, ont sans doute plus alimenté les débats philosophiques que l'inverse. Ses définitions du concept d'entropie ont définitivement « assimilé » les probabilités, malgré plusieurs questions ouvertes quant à leur interprétation, et, sans doute pour cette même raison, ont atteint un niveau d'abstraction propice aux plus diverses interprétations et aux multiples récupérations dans des contextes forts différents. Le principe entropique perdait son caractère absolu : « l'impossibilité d'une diminution non compensée [pour un système isolé] d'entropie semble être réduite à une improbabilité »⁶⁰ (Gibbs 1876 : 229).

⁶⁰ « [...] *the impossibility of an uncompensated decrease of entropy seems to be reduced to improbability.* »

3.3.4 Aparté : Planck et Einstein

Max Planck (1858-1947) et Albert Einstein (1879-1955) n'ont plus besoin de présentation. Si le second a une réputation qui dépasse largement le domaine de la physique, il la doit en partie au premier, dont les travaux lui ont servi d'inspiration et qui a eu, dans le contexte de l'époque, le flair et l'audace d'endosser ses premiers articles. Les deux hommes s'admiraient beaucoup⁶¹. Ils ont travaillé tous les deux sur la thermodynamique et la mécanique statistique aux prémices de la mécanique quantique.

Les premiers travaux théoriques de Planck ont touché la thermodynamique, son « premier amour » (Kuhn 1978 : 12), et plus particulièrement le concept d'entropie⁶². On peut même affirmer qu'il a donné à ce concept sa « formulation classique définitive » (Uffink 2001 : 342). Selon lui, le second principe de la thermodynamique est indispensable à la détermination des processus et plus spécifiquement leur directionnalité. Il s'agirait d'un principe empirique dont la preuve ne peut être déduite que d'une « loi de l'expérience qui ne ferait aucun doute » (Planck 1945 : §116), soit l'impossibilité du mouvement perpétuel de second type. Il en découle le principe entropique, comme quoi l'entropie totale, celle d'un système quelconque *et* de son environnement, ne peut qu'augmenter. Celle-ci demeure constante que dans un processus réversible, qui demeure un « cas limite idéalisé » (Planck 1945 : §115). Il affirme, par ailleurs, que la signification physique du second principe, qui doit rendre compte de l'irréversibilité, ne peut être ramenée aux métaphores de la « dégradation de l'énergie » ou de la « perte de travail », car il existe des phénomènes irréversibles, comme le mélange de deux gaz parfaits, dont les états initial et final présentent la même forme d'énergie.

Mais Planck a aussi contribué aux développements de la version statistique du concept d'entropie. Il s'est intéressé de très près dans les années 1890 à la théorie de la lumière et à celle de la radiation thermique appliquées aux corps se rapprochant des « corps noirs », cas idéalisés de corps absorbant toute la radiation qu'ils reçoivent. La fréquence à laquelle émettent les corps, en général et en particulier les corps noirs, c'est-à-dire leur couleur, dépend de leur température. Les théories d'alors peinaient à rendre compte de manière satisfaisante des résultats expérimentaux quant à la quantité, ou densité, de radiations émises en fonction de la fréquence et la température d'émission. L'idée de Planck est alors d'établir une « entropie de radiation », par analogie à l'énergie de radiation, qu'il présenta à

⁶¹ Une admiration parfois teintée de petite malice. Après la mort de Planck et plusieurs années après une expérience importante corroborant la théorie de la relativité générale, Einstein affirma à son sujet : « *He was one of the finest people I have ever known [...] but he didn't really understand physics, [because] during the eclipse of 1919 he stayed up all night to see if it would confirm the bending of light by the gravitational field. If he had really understood [the general theory of relativity], he would have gone to bed the way I did.* » (Einstein, in Coles 2000 : 53)

⁶² Son *Vorlesungen über Thermodynamik* a connu onze éditions entre 1897 et 1966 et a fait autorité dans la première moitié du XX^e siècle.

Boltzmann, qui répondit qu'une théorie statistique de la radiation devait inclure un élément de discontinuité jusqu'alors inconnu (Nugayev 2000). Au terme du travail « le plus ardu de [sa] vie » (Planck, in Kuhn 1978 : 97), il obtint une formule plus juste que ses concurrentes. Il se lance alors dans la « tâche de trouver sa réelle signification physique », ce qui l'a mené à considérer la « relation entre entropie et probabilité, et donc au raisonnement de Boltzmann » (Planck, in Kuhn 1978 : 98). Ce raisonnement le conduisit, dans un « coup mathématique habile » (Cropper 2001 : 234), à l'entropie combinatoire de Boltzmann, qu'il applique aux corps noirs.

Albert Einstein (1879-1955) s'est inspiré de Planck pour ses travaux sur la mécanique statistique. Lorsqu'il fût engagé à l'Académie des sciences à Berlin, le chimiste Walter Nernst le baptisa « *Boltzmann redivivus* » (Renn 2000 : 7), car son embauche était liée surtout à ses travaux sur l'atomisme et la mécanique statistique, des sujets de prédilection de Boltzmann. En effet, Einstein a manifesté une préoccupation constante pour ces sujets, y dédiant une quarantaine d'articles, dont son tout premier en 1900. De plus, selon Klein (1967 : 512), les idées les plus originales d'Einstein sont toutes « intimement liées à sa compréhension de la thermodynamique ». Bien que ses écrits n'aient pas contribué de manière significative aux fondements de la version statistique du second principe de la thermodynamique, il a su en tirer profit par un raisonnement s'appuyant sur des analogies formelles. Gingras (2011 : 94) affirme :

Même s'il ne nous a laissé aucune trace des calculs qui aboutirent au fameux article expliquant l'effet photoélectrique, il est probable qu'Einstein commença par manipuler l'équation de Planck sans en tirer une intuition probante à l'aide de l'approche entropique, alors qu'il y parvint ensuite avec l'approximation de Wien, le calcul de l'entropie le conduisant à cette formule élégante [...] pour la variation de l'entropie du rayonnement contenu dans [un volume donné].

Ce ne fut pas automatique, car ses travaux se situaient dans une « transition entre une théorie physique située au cœur de la conception mécaniste du monde, ancrée dans le XIX^e siècle, et une théorie moderne » (Barberousse 2002 : 196).

Einstein a écrit trois articles traitant de la mécanique statistique de 1902 à 1904. Ses connaissances des écrits de Boltzmann étaient alors, semble-t-il, fragmentaires et il n'était pas du tout au courant de ceux de Gibbs. Le premier article aborde les définitions de la température et de l'entropie dans des conditions d'équilibre thermique et avec le théorème d'équipartition ; le second traite de l'irréversibilité ; et le troisième des fluctuations et de nouvelles méthodes pour déterminer la constante de Boltzmann. Il s'était convaincu très tôt de la justesse de l'hypothèse atomique, au moins sous la forme minimale selon laquelle les systèmes mécaniques possèdent un nombre fini de degrés de liberté. Son but était de dériver les lois de la thermodynamique à partir des hypothèses physiques les plus

générales et du calcul des probabilités, en évitant, tout comme Gibbs, toute hypothèse spécifique sur la nature des constituants élémentaires des systèmes physiques ou de leurs interactions. Il appert qu'il suivait le chemin tracé par Boltzmann :

Les concepts de température et d'entropie découlent du principe de conservation de l'énergie et la thèse atomiste, de même que le second principe dans sa forme la plus générale, soit l'impossibilité du mouvement perpétuel de second type, si l'on suppose que les distributions d'état des systèmes isolés n'évoluent jamais vers des états plus improbables.⁶³

⁶³ « *The concepts of temperature and entropy follow from the assumption of the energy principle and the atomistic theory, and so does also the second law in its most general form, namely the impossibility of a perpetuum mobile of the second kind, if one uses the hypothesis that state distributions of isolated systems never evolve into more improbable ones.* » (Einstein 1903, in Navarro 1998 : 160-1)

3.4 Autres disciplines

L'abstraction du concept d'entropie issue du formalisme de la mécanique statistique a ouvert la voie à de multiples récupérations par analogies formelles. Ainsi, tout ce qui se laisse décrire par une relation entre différents niveaux d'un même système, où l'un de ces niveaux est sous-déterminé, semble une occasion de parler d'entropie ; il en va ainsi de l'« entropie politique du choix de vote » (Gill 1997). Cette sous-détermination n'est pas sans rappeler la notion d'information, plus précisément un manque d'information, qu'un soi-disant « démon » pourrait combler. Il n'en faut pas davantage pour relier l'entropie à l'incertitude, ce qu'a fait Shannon. La liste pourrait encore être longue, d'autant plus qu'elle continue de s'écrire. Au lieu de déplorer un tel foisonnement qui frôle la cacophonie, il faut se réjouir de la richesse du concept d'entropie et des connexions interdisciplinaires qu'il suggère et entretient. Mais, en tout état de cause, un examen rigoureux de sa signification, de ses interprétations, allant au-delà des analogies, est nécessaire et même urgent (Chapitres 4 et 5).

3.4.1 Théorie de l'information

La connexion entre le concept d'entropie et celui d'information remonte, on l'a vue, à l'expérience de pensée de Maxwell. L'information, qui demeure un concept difficile à définir clairement, possède intuitivement des propriétés quantitatives, comme sa capacité à être « encodée » ou encore sa « non-négativité ». Ce type de propriétés est l'objet de la *théorie mathématique de l'information*, dont l'origine remonte à l'œuvre majeure de Claude Shannon (1916-2001), publiée dans le *Bell System Technical Journal* en 1948 sous le titre « Une théorie mathématique de la communication »⁶⁴.

Il souligne d'entrée de jeu que le « problème fondamental de la communication est de reproduire en un point, soit exactement soit approximativement, un message sélectionné en un autre point » (Shannon 1949 : 1). En théorie générale de l'information, trois concepts clés se partagent l'essentiel de son contenu, soit les concepts d'émetteur, de récepteur et de message. La théorie de Shannon considère des ensembles de messages discrets, formés d'une séquence de symboles prédéterminés, pour lesquels des valeurs de probabilité p_i sont attribuées. L'idée de base est de caractériser la quantité d'information par une fonction, H , ne dépendant que de la probabilité des messages et respectant un ensemble d'axiomes. Puisque les probabilités sont multiplicatives et que l'information est additive, comme quoi le

⁶⁴ « *A Mathematical Theory of Communication* ».

gain en information de deux communications indépendantes est la somme de l'information des communications, mais aussi continue, la fonction H est nécessairement logarithmique :

$$H(p_i) = -K \sum p_i \log p_i.$$

où K est une constante et le logarithme est en base 2. Cette fonction est généralement interprétée comme une « mesure du manque d'information », ou comme « mesure d'incertitude » ou « d'ignorance », pour une situation communicationnelle donnée. Elle présente ainsi quelques caractéristiques idoines à certaines intuitions quant à l'information. Par exemple, H est nulle lorsque toutes les probabilités p_i sont nulles sauf une étant égale à l'unité, ce qui équivaut à une situation de certitude. Ou encore, H est maximale lorsque toutes les probabilités p_i sont égales, ce qui correspond à la situation la plus incertaine.

La fonction $H(p_i)$ permet de caractériser quantitativement l'information et la traite, d'une certaine façon, tel un « phénomène physique » (Floridi 2004 : 53). Elle n'offre toutefois pas d'éclairage pertinent sur les divers sens du terme « information », sur la connaissance en général, et elle est étrangère à la signification d'un message individuel. Shannon (1948 : 11) note que la « forme de H peut être reconnue comme celle de l'entropie telle que définie dans certaines formulations de la mécanique statistique ». L'origine de cette analogie formelle mérite sa place dans le registre des anecdotes savoureuses de l'histoire des sciences.

Mon plus grand souci était de comment l'appeler. J'ai pensé l'appeler 'information', mais le mot était galvaudé, en sorte que j'ai décidé de l'appeler 'incertitude'. Lorsque j'en ai discuté avec John von Neumann, il avait une meilleure idée. Von Neumann m'a dit : 'tu devrais l'appeler entropie, et ce pour deux raisons. D'abord, ta fonction d'incertitude a été utilisée en mécanique statistique sous le même nom, elle a donc déjà un nom. Ensuite, et surtout, personne ne sait vraiment ce qu'est l'entropie, de sorte que dans un débat tu auras toujours l'avantage'.⁶⁵

L'idée de von Neumann n'est peut-être pas aussi heureuse que ne le pensait Shannon, car elle a davantage obscurci les concepts d'entropie et d'information que l'inverse (Maroney 2005 ; Müller 2007 ; Ben-Naim 2008). Mais le lien est néanmoins substantiel. Car, on l'a vu, les expériences de pensée comme celle du « démon de Maxwell », malgré les confusions qui l'entourent, révèlent la force de ce lien même si un long travail d'éclaircissement est nécessaire (Chapitre 5). On a vu aussi que l'entropie statistique implique bien sûr les probabilités, mais en outre, dans une certaine mesure, la

⁶⁵ *My greatest concern was what to call it. I thought of calling it 'information', but the word was overly used, so I decided to call it 'uncertainty'. When I discussed it with John von Neumann, he had a better idea. Von Neumann told me, 'you should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, nobody knows what entropy really is, so in a debate you always have the advantage'.* (Shannon, in Tribus & McIrvine 1971).

théorie ergodique, alors même que la théorie mathématique de l'information peut être considérée comme une branche de la théorie des probabilités, et même de la théorie ergodique (Gray 2011), bien qu'il semble juste de dire qu'elle s'en soit partiellement affranchie (Klir 2006).

3.4.2 Mécanique quantique

L'entropie quantique est bien entendu postérieure au développement de la mécanique quantique dans les années 1920 et elle est due à John von Neumann (1955). Son argumentation comporte deux étapes et s'appuie sur une expérience de pensée semblable, de l'avis de son auteur, au « démon de Maxwell », où un système est composé d'une enceinte contenant les molécules d'un gaz parfait et des membranes semi-perméables permettant d'effectuer des tests quantiques, soit de séparer les molécules selon leur état quantique. L'argumentation de von Neumann est particulière en ceci qu'elle fait intervenir autant la mécanique quantique que la thermodynamique au sein de la même expérience de pensée. Comme aucune de ces théories n'est déductible l'une de l'autre, il ne peut offrir une déduction formelle. Il s'appuie plutôt sur une expérience de pensée soutenant elle-même une hypothèse. Cette hypothèse est alors appliquée dans un argument arithmétique où une certaine quantité, soumise au formalisme quantique, se comporte de manière *analogue* à l'entropie thermodynamique lors d'un processus illustré par cette expérience de pensée.

L'argumentation présuppose la validité des deux principes fondamentaux de la « thermodynamique phénoménologique », soit l'« impossibilité du mouvement perpétuel de premier et second type » (von Neumann 1955 : 359). Dans la première étape de l'argumentation, il soutient que les *mesures* quantiques *augmentent* l'entropie thermodynamique du système mesuré. Selon l'intuition fournie par la thermodynamique, cela peut signifier que le système mesuré reçoit une quantité de chaleur, peut-être en raison d'un investissement d'une quantité de travail dans l'appareil de mesure. Dans la seconde étape, il cherche un équivalent mécanique quantique qui se comporte comme l'entropie. Dans une étape du processus cyclique décrit par l'expérience de pensée, le gaz est comprimé isothermiquement, ce qui demande, d'après un calcul standard de la thermodynamique appliqué à un gaz parfait, une quantité de travail égale à $NkT \ln 2$, compensée par une quantité de chaleur équivalente, et donc l'entropie du gaz diminuerait de $(NkT \ln 2)/T = Nk \ln 2$. Comme cette diminution d'entropie doit être compensée, von Neumann en conclut que l'entropie doit avoir augmenté quelque part. Il fait alors intervenir son hypothèse et soutient que l'augmentation se produit lors de la mesure dans l'une des étapes du cycle. Enfin, von Neumann identifie une quantité présentant le comportement recherché :

puisque $S_{VN} \equiv -k_B \text{Tr}[\rho \ln \rho]$ augmente lors de la mesure selon le formalisme de la mécanique quantique, il avance qu'il s'agit du « corrélat mécanique-quantique de l'entropie thermodynamique ».

Compte tenu de l'argumentation décrite précédemment, il est pour le moins surprenant que S_{VN} soit aussi bien accepté comme l'entropie quantique. Selon Shenker (1999), l'argument de von Neumann serait inopérant car l'idée que l'entropie des états purs est nulle n'est pas appliquée de manière consistante, en sorte que sa conclusion ne découle pas de ses prémisses. De plus, si S_{VN} est bien l'analogue quantique de l'entropie, alors les partitions semi-perméables se révèlent être ni plus ni moins que des machines à mouvement perpétuel. En fait, les débats quant à l'entropie quantique risquent de demeurer stériles tant et aussi longtemps qu'une question fondamentale ne sera pas réglée, à savoir celle du comportement de l'entropie lors d'un processus de mesure, si elle augmente ou si plutôt elle diminue. Certains ont avancé l'idée que la mesure quantique produit une *augmentation* de l'entropie du système mesuré car celui-ci devient alors « couplé » à l'appareil de mesure qui possède plusieurs degrés de liberté (Jancel 1963 ; Partovi 1989). D'autres, à l'inverse, ont soutenu qu'il y a *diminution* de l'entropie car la mesure fournit de l'information et que ce gain d'information exige un rejet d'entropie quelque part (Lubkin 1987).

On constate que ces considérations, toujours d'actualité, sont au cœur de plusieurs préoccupations épistémologiques importantes, comme les expériences de pensée, le fameux problème de la mesure et le lien entre entropie et information (Jodoin 2010b ; discuté au Chapitre 5).

3.4.3 Biologie

Historiquement, mais aussi conceptuellement, la récupération du concept d'entropie en biologie se présente sous deux aspects : (i) selon une approche *substantielle*, en ce sens que l'entropie thermodynamique est simplement appliquée à une classe de phénomènes particuliers, les êtres vivants, considérés comme des systèmes thermodynamiques (généralement) ouverts ; et (ii) selon une approche *analogique*, où une fonction d'un espace des phases (généralement) génotypique trouve des similitudes avec l'entropie statistique et s'insère conséquemment dans le formalisme de la mécanique statistique (Chapitres 6)⁶⁶.

La relation entre la mécanique statistique et la biologie de l'évolution a une longue histoire (voir Provine 1971 ; Morrison 1997 et 2007 ; Jodoin 2014b). L'utilisation des méthodes statistiques dans la

⁶⁶ Malheureusement, l'emploi, sur les modes analogique et substantiel, du concept d'entropie en biologie ne peut recevoir ici un traitement exhaustif pour des raisons d'espace.

caractérisation du comportement des grands ensembles idéalisés constitue la base de l'*analogie* entre les deux disciplines. Par exemple, l'un des objectifs de Ronald A. Fisher (1890-1962) était de combler les lacunes explicatives du modèle mendélien qui pouvait prédire la descendance d'un individu mais pas d'une population. Il a ainsi été le premier à fournir une véritable description des opérations de la sélection comme « un processus de remplacement des gènes dans une population mendélienne » (Morrison 2007 : 319). L'énoncé de ce que Fisher (1930 : 35) appelle le *théorème fondamental de la sélection naturelle* (TFSN) est le suivant : « le taux d'accroissement de la fitness de tout organisme à n'importe quel moment est égal à sa variance génétique en fitness à ce moment-là ». Formellement, $dM/dt = \mathcal{W}$, où M est la fitness moyenne de la population et \mathcal{W} est ce que Fisher appelle la « variance génétique en fitness » (« *genetic variance in fitness* »). Puisque la variance \mathcal{W} est une quantité non négative, il s'ensuit que la fitness populationnelle moyenne M est une quantité non décroissante.

Selon Fisher, le TFSN suggère ainsi que les changements évolutifs attribuables à la sélection naturelle sont intrinsèquement directionnels, en ce sens qu'ils évolueraient vers des états finaux caractéristiques. Ce théorème présentait selon lui une « ressemblance remarquable avec la seconde loi de la thermodynamique » (Fisher 1930 : 36). Dans les deux cas, il s'agit de propriétés de populations qui demeurent « vraies » peu importe la nature des éléments qui les composent et dans les deux cas il s'agit de « lois statistiques » requérant un accroissement continu d'une quantité « mesurable », en l'occurrence l'entropie ou la fitness. À l'instar de la seconde loi qui stipule qu'un système isolé évolue toujours d'un état à un autre d'entropie plus élevée, le TFSN semble indiquer qu'une population évolue toujours, sous l'action de la sélection naturelle, d'un état possédant une fitness relativement basse à un état de fitness plus élevée. Ainsi, la sélection naturelle serait un processus de changement dirigé. (Pour un examen critique de cette récupération conceptuelle, voir Jodoin 2013d et 2014b).

Outre cette récupération conceptuelle par analogie formelle, le concept d'entropie est aussi utilisé en biologie selon une approche substantielle, c'est-à-dire par une application de la physique aux systèmes biologiques. Déjà, on l'a vu, Thomson (1852, in 1882 : 514) tirait des conclusions du second principe à l'effet que la terre deviendra tôt ou tard « inhabitable pour l'homme ». Le caractère apparemment inéluctable de l'augmentation de l'entropie a aussi donné l'impression qu'il était incompatible avec le phénomène de la vie. Mais une compréhension élémentaire du principe entropique (Chapitre 4) est suffisante pour balayer cette impression. En effet, si l'entropie d'un être vivant peut diminuer, ce qui équivaut, par abus de langage, à ce qu'on a appelé un apport de « néguentropie », alors nécessairement celle de son environnement augmente, ce qui n'est pas fondamentalement différent, du point de vue physique du moins, du fonctionnement d'un réfrigérateur, un système ouvert et un processus « forcé » selon la distinction de Clausius.

En terminant, et pour bien montrer toute la richesse mais aussi, en quelques sortes, l'amphibologie du concept d'entropie, voici trois autres exemples. D'abord, dans un ouvrage désormais célèbre, issu d'une série de conférences données à Dublin, Schrödinger (1944) a évoqué la nature « énigmatique » des organismes vivants qui semblent « éviter la décadence vers l'équilibre »⁶⁷ en se nourrissant d'« *entropie négative* ». S'appuyant sur la formulation de l'entropie de Boltzmann, il interprète alors l'entropie comme une mesure du désordre, en sorte que la capacité d'un organisme vivant à se maintenir dans un état stationnaire d'ordre relativement élevé consiste à « tirer continuellement de l'ordre de son environnement »⁶⁸. Ensuite, l'hypothèse Gaïa laisse entendre que la Terre est un organisme luttant contre des forces entropiques l'amenant vers le désordre (Wikipédia, « *Gaïa Hypothesis* »⁶⁹). Enfin, Chalmers (1996) a proposé l'idée que l'entropie au sens de Shannon était la clé pour comprendre rien de moins que le dualisme corps-esprit.

⁶⁷ « *evade the decay to equilibrium* ».

⁶⁸ « *continually sucking orderliness from its environment* ».

⁶⁹ Je me permets ici de citer l'encyclopédie en ligne *Wikipédia*, car compte tenu de la nature de l'hypothèse il me paraît approprié d'avoir recours à ce type de référence.

3.5 Conclusion

Au terme de cet historique, aux allures d'une épopée, parfois « tragicomique » pour reprendre l'expression de Truesdell (1980) à propos de l'histoire de la thermodynamique, on constate à quel point le concept d'entropie joue un rôle majeur dans plusieurs disciplines scientifiques. On constate par ailleurs que le concept d'entropie ne se laisse pas facilement appréhender. Il fut d'abord motivé par un effort technologique d'améliorer le fonctionnement des machines à vapeur, et donc issu des développements de la thermodynamique, avec les travaux de Carnot, Clausius et Thomson, dans un contexte de révolution industrielle. Il découle du second principe de la thermodynamique, selon lequel il ne peut y avoir conversion complète de la chaleur en travail. Curieusement, ce principe a d'abord été proposé, en substance quoiqu'indirectement, par Carnot, avant même que l'énergie soit un concept bien établi. Il semble en effet que les définitions opératoires des transferts d'énergie, sous forme de chaleur ou de travail, étaient suffisamment claires pour qu'on puisse avancer théoriquement, mais modestement, sans l'unification conceptuelle offerte par le concept d'énergie.

L'apport majeur de Carnot est d'avoir offert une quantification de la conversion de la chaleur, le « calorique », en travail, la « puissance motrice ». Et ce faisant il offrait aussi une expression quantitative simple, dépendant seulement des températures d'opération, de l'impossibilité du mouvement perpétuel. Thomson a ensuite reformulé ces résultats, à la lumière des travaux de Joule, et ainsi donné les expressions du premier et du second principe de la thermodynamique, alors naissante. Mais c'est à Clausius que l'on doit le concept original de l'entropie. Il a ainsi exprimé par une fonction d'état ces « valeurs d'équivalence » en supposant un cas limite idéal exemplifiant, comme chez Carnot, la réversibilité, où ces valeurs se compensent mutuellement. Cette fonction d'état a été désignée par la lettre S , pour aucune raison apparente, et nommée « entropie » à partir du grec pour « transformation » et du préfixe « en » pour sa proximité avec le mot « énergie ». Mais cette fonction d'entropie demeure indéterminée pour un processus irréversible, car elle n'offre qu'un seuil minimal pour un cas de réversibilité, ni pour un système qui n'est pas à l'équilibre. Les travaux subséquents, toujours en cours de développement, en thermodynamique traitant du non-équilibre et des processus irréversibles n'ont pas modifié substantiellement le concept d'entropie.

À la fin du XIX^e siècle, il s'est retrouvé au cœur de débats portant sur les constituants de la matière et les fondements de la physique, avec les travaux de Maxwell et Boltzmann. Il a par la suite été récupéré, sans doute en raison de son côté abstrait, dans plusieurs disciplines, comme la mécanique quantique, la théorie des communications et la génétique des populations. En ce sens, il s'agit bien là d'un concept « protéiforme » (Balian 2003). Et sa difficile appréhension exprime certaines difficultés

profondes de l'épistémologie. En effet, d'une part il découle de l'observation de phénomènes physiques maintes fois répétés comme la diffusion de la chaleur ou le fonctionnement de machines et dont le traitement théorique au moyen de ce concept semble parfaitement corroboré. D'autre part, et c'est là le principal schisme conceptuel ouvrant sur la mécanique statistique, il reçoit une formulation et une interprétation à partir d'un traitement mathématique à partir d'une théorie jugée fondamentale, la mécanique, appliquée à un grand nombre d'entités inobservables, d'où les statistiques, mais considérées comme les constituants de la matière observable. L'opposition entre empirisme et rationalisme trouve donc ici l'une de ses expressions les plus vives.

Les écrits de Maxwell et de Boltzmann, les articles comme les correspondances, indiquent bien la nature des débats qui animaient leurs travaux en mécanique statistique. Le premier n'a d'ailleurs jamais tenté d'offrir une interprétation mécaniste du second principe ou de l'entropie, lesquels étaient, selon lui, des concepts de nature statistique, ce que tente d'illustrer son expérience de pensée, le « démon de Maxwell ». Cette expérience, qui fait encore l'objet de débats, laisse entendre que l'entropie « survient » en raison de nos limitations sensorielles et cognitives à suivre la trajectoire d'un très grand nombre de molécules. Le second a suivi les traces du premier, en ceci qu'il s'est appuyé sur l'hypothèse du *Stoßzahlansatz* et sur une distribution stationnaire des vitesses des nombreuses molécules constituant un gaz macroscopique, c'est-à-dire à notre échelle, pour représenter un l'état d'équilibre. L'accomplissement majeur de Boltzmann est ici d'avoir identifié une quantité, fonction de variables mécaniques, qui décroît dans le temps avant de se maintenir fixe lorsque la distribution des vitesses est de type maxwellien, ce qui reflèterait l'accroissement (inverse) de l'entropie qui atteint un maximum à l'équilibre.

Mais la version statistique de l'entropie a reçu un autre traitement avec la méthodologie des ensembles développée par Gibbs. Très formelle et rigoureuse, à l'image de son fondateur, cette approche se laisse cependant difficilement appréhendée par l'interprétation, procédant pour l'essentiel par analogie. Une méthode qui semble avoir fait école puisqu'Einstein a aussi procédé de la sorte pour l'explication de l'effet photoélectrique, de même Shannon pour la quantification de l'information, von Neumann pour la version quantique de l'entropie et, dans une certaine mesure, Fisher en génétique des populations.

4 Critique du concept d'entropie I

RÉSUMÉ. Le concept d'entropie fait paradoxalement consensus sur deux points : il est fondamental et personne ne sait vraiment de quoi il s'agit. Il importe donc d'apporter un examen épistémologique critique de ce concept afin d'en définir le sens et d'en identifier le caractère fondamental ainsi que les usages légitimes selon les contextes explicatifs pertinents. Le présent chapitre focalise sur la version *thermodynamique* du concept d'entropie : discussions sur l'irréversibilité (« flèche du temps »), définitions et principes fondamentaux de la thermodynamique, signification du second principe comme étant distinct du principe entropique, construction et propriétés de la fonction d'état définissant l'entropie, la question du non-équilibre, et la valeur des stratégies explicatives impliquant le principe entropique.

You can't win with a loosing hands

Bob Dylan

Paradoxalement, on l'a vu, le concept d'entropie fait consensus sur deux points : il est fondamental et personne ne sait vraiment de quoi il s'agit. Comme le soulignait Khinchin déjà en 1949 (: 137), très peu de notions scientifiques ont reçu autant d'efforts afin de clarifier leur signification théorique et philosophique.

Plusieurs auteurs ont donné leur imprimatur au second principe de la thermodynamique en le présentant comme l'un des principes les plus fondamentaux de la physique moderne, voire le plus fondamental. Incidemment le concept d'entropie censé exprimer le sens de ce principe serait un concept physique essentiel. Ainsi, par exemple, Planck (1945 : 107) était persuadé que les futurs métaphysiciens considéreraient le principe entropique (version Planck) comme rien de moins qu'une « vérité apriorique ». Eddington (1928 : 74-5) a écrit ces mots maintes fois cités, qu'il vaut la peine de présenter ici, exprimant le statut privilégié de ce principe :

La loi selon laquelle l'entropie augmente toujours — la seconde loi de la thermodynamique — occupe, je crois, la position suprême parmi les lois de la nature. Si quelqu'un vous fait remarquer que votre théorie fétiche de l'univers est incompatible avec les équations de Maxwell — alors tant pis pour les équations de Maxwell. Si elle se trouve réfutée par l'expérience — eh bien, ces expérimentateurs connaissent aussi leurs

ratées. Mais si votre théorie est en contradiction avec la seconde loi de la thermodynamique, je ne vous accorde aucun espoir ; il ne vous reste qu'à vous écrouler dans la plus profonde humiliation.¹

Selon Einstein (1951), la thermodynamique classique est la seule théorie dont le cadre théorique fondamental ne serait jamais révolu. Enfin, et la liste pourrait être encore longue, la validité du second principe de la thermodynamique, pour Brillouin (1949), serait aussi assurée que les lois fondamentales de la mécanique classique.

Le concept d'entropie est pourtant mal compris et souffre d'une multiplicité définitionnelle, d'une polysémie. Ainsi Capek & Sheehan (2005) ne présentent pas moins de vingt-et-une formulations de l'entropie et évoquent les travaux de Hänggi qui en retracerait cinquante-cinq variétés. L'entropie serait ainsi un concept étrange, difficile et multiforme (Darrigol 2003) ; bref, un concept protéiforme (Balian 2003) ; énigmatique (Capek & Sheehan 2005) ; abstrus (Atkins 1987) ; dont le consensus quant à sa signification reste inaccessible (Grandy 2008) ; différent de la plupart des autres quantités physiques (Wherl 1978) ; voire un concept mal nommé et ne signifiant, à toutes fins pratiques, rien (Ben-Naim 2008). Mais il serait (aussi) le concept le plus influent de la mécanique statistique (Sethna 2011) ; et le plus fascinant de la thermodynamique (Rigden 2005) ; un concept nécessaire à la compréhension du comportement des systèmes macroscopiques (Gould & Tobochnik 2006) ; utile à l'analyse des dispositifs d'ingénierie (Çengel & Boles 2006) ; représentant une quantité physique fondamentale (Lieb & Yngvason 2003 ; Brissaud 2005) ; un concept qui mérite le respect (Atkins 1987) ; et plus fondamental que le concept de température (Lieb & Yngvason 2003 ; Thess 2011).

Cette panoplie de références ne fait qu'illustrer l'effervescence entourant ce concept. Et pourtant elle n'est qu'un aperçu. Il a été vu (Chapitre 3) que l'origine du concept d'entropie remonte à la révolution industrielle alors que les ingénieurs tentaient de comprendre et d'optimiser le rendement des machines à vapeur. Les travaux de Sadi Carnot, qui s'appuient sur l'impossibilité du mouvement perpétuel, ont constitué un élément majeur dans l'heuristique du concept d'entropie en ce qu'ils présentent le concept de cycle réversible duquel une quantité maximale de travail peut être extraite d'une quantité de chaleur donnée. Mais ils montrent surtout que la conversion de la chaleur en travail est limitée de principe, ce qui signifie que le rendement de ces machines à vapeur est aussi limité. La formalisation mathématique ainsi que le nom de l'entropie (du grec « transformation ») sont dues au physicien allemand Rudolf Clausius. L'idée générale derrière cette formalisation est qu'une quantité – $\delta Q/T$ – se conserve lors d'un processus réversible entre deux états d'équilibre mais ne se conserve par

¹ « *The law that entropy always increases — the second law of thermodynamics — holds, I think, the supreme position among the laws of Nature. If someone points out to you that your pet theory of the universe is in disagreement with Maxwell's equations — then so much the worse for Maxwell's equations. If it is found to be contradicted by observation — well, these experimentalists bungle things sometimes. But if your theory is found to be against the second law of thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation.* »

lors d'un processus irréversible. Comme la tendance à l'équilibre des systèmes isolés de leur environnement constituait déjà un fait physique bien établi (Fourier), cette quantité est censée augmenter puis se maintenir dans le temps pour tout système isolé de son environnement et atteignant l'équilibre après un certain temps. Or, un processus réversible est une idéalisation qui n'est donc jamais instanciée. Tous les processus ne sont donc pas réversibles, ils sont donc irréversibles. Le concept d'entropie est ainsi articulé au sein d'un principe, le principe entropique, stipulant l'augmentation de l'entropie dans le temps (à cette nuance près qu'une diminution locale est possible, discuté plus loin). Ce principe est donc censé fournir une explication des processus vers l'équilibre et donc irréversibles *parce que l'entropie doit augmenter*.

Les concepts de quantité non-conservative² et de processus irréversible semblaient incompatibles à première vue avec la mécanique newtonienne, qui est réversible par rapport au renversement du temps. Ludwig Boltzmann a ainsi tenté de traduire ce concept en termes de quantités statistiques dérivées des lois mécaniques. Cette reformulation a permis d'étendre l'application du concept d'entropie aux systèmes qui ne sont pas à l'équilibre bien qu'elle repose sur un certains nombres d'hypothèses discutables. Elle a par ailleurs ouvert la voie, sans doute en raison de son caractère abstrait, à maintes interprétations et à des récupérations par analogie formelle dans d'autres disciplines scientifiques, comme la génétique des populations, la mécanique quantique ainsi que la théorie de l'information. Plusieurs questions demeurent donc ouvertes : quel est la bonne définition du concept d'entropie, quel sens a-t-il et quelle est son articulation au sein du second principe, à quels types de systèmes physiques réfère-t-il, comment l'interpréter, et quelle est sa portée dans les explications scientifiques ? La caractéristique principale du concept d'entropie, du moins celle qui soulève le plus débats, est sans nul doute l'irréversibilité. Puisque le principe entropique est jugé fondamental et qu'il est pour ainsi dire le seul à manifester l'irréversibilité, il est mobilisé au sein d'explication de « phénomènes » irréversibles tout aussi fondamentaux. Il expliquerait ainsi rien de moins que la direction du temps (Reichenbach 1956), ou encore la directionnalité de la sélection naturelle (Rosenberg 2011).

Ce chapitre s'ouvre donc sur une discussion sur l'irréversibilité (section 2). Deux sens sont ici distingués, soit celui de la non-invariance par renversement du temps et celui de l'irrécupérabilité. Face à cette apparente opposition avec la réversibilité des lois physiques, il est opportun de s'interroger sur l'origine de l'asymétrie thermodynamique et de l'irréversibilité ainsi définies par l'augmentation

² Une *force* est conservative si elle dépend d'un potentiel, qui lui ne dépend pas du chemin d'intégration, autrement dit du chemin suivi par son point d'action. De sorte que le travail effectué par cette force est aussi indépendant du chemin d'intégration et ne dépend que du potentiel. Une *quantité* est dite conservative s'il est possible de construire une équation de conservation de sorte qu'elle se conserve dans le temps. C'est le cas notamment de l'énergie mécanique.

d'entropie. L'approche de Boltzmann de la mécanique statistique offre une réponse en ceci que la croissance d'entropie exprimerait la tendance moyenne d'un système, composée de plusieurs molécules, d'évoluer vers des états de plus en plus probables selon une caractérisation de la gamme des états microscopiques compatibles avec les paramètres macroscopiques de ce système. Mais il n'y a pour ainsi dire rien de consensuel dans cette réponse (Chapitre 5). De plus, le concept d'entropie est censé être au centre de stratégies visant à expliquer plusieurs phénomènes irréversibles, comme les transferts thermiques, un œuf se brisant au sol, le vieillissement, ou de la direction même du temps. Il s'agit de processus qui soit ne peuvent être effectués à l'envers, soit ne retrouvent jamais leur état initial. Cette prétention associée à l'entropie s'est amplifiée au point de rejoindre les tentatives d'explication de la direction du temps lui-même, qui va inexorablement du passé vers l'avenir, mais aussi de plusieurs asymétries temporelles, comme quoi nous avons des souvenirs du passé mais pas de l'avenir, que nous pouvons influencer le futur mais pas le passé, etc. Le principe entropique se présente donc autant comme l'*explanandum* d'une asymétrie temporelle que comme son *explanans*. C'est pourquoi ce principe est censé rendre compte de la « flèche du temps ». Il y a cependant beaucoup de confusion autour de cette notion, dont, au premier chef, la notion de temps elle-même. Elle devra donc subir un examen critique afin de circonscrire adéquatement ce qui influence les interprétations du concept d'entropie et sa portée dans les explications en général.

La troisième section porte sur la thermodynamique et plus précisément sur le concept d'entropie au sein de cette discipline, laquelle présente plusieurs déclinaisons. Suite à une présentation de ce que représente un état et un processus en thermodynamique, le second principe est discuté. Or il a été vu que celui-ci pouvait recevoir plusieurs formulations. La discussion porte alors sur une reconstruction de ce principe à partir de quelques fondamentaux jugés suffisants pour établir le « noyau dur » de sa signification. En termes simples, je soutiens que cette signification réfère à l'impossibilité de convertir complètement une quantité de chaleur en travail pour un système donné. Elle suppose donc une définition claire des concepts de chaleur et de travail. Cette reconstruction, qui ne s'éloigne pas significativement des travaux fondateurs de la discipline mais plutôt s'en inspire, permet ainsi de distinguer le second principe du principe entropique. En effet, les problèmes épistémologiques que rencontrent ces principes ne sont pas nécessairement coextensifs car il font intervenir des concepts différents qui reçoivent aussi des significations de théories différentes (par ex. il n'y a pas de concept statistique du travail). Afin de bien cerner la stratégie explicative du principe entropique (section 3.7), une présentation de l'argumentation visant à définir la fonction d'état correspondant à l'entropie suit (section 3.4.1), ainsi qu'une présentation de ses propriétés (section 3.4.2). Le principe entropique, qui stipule que l'entropie d'un système isolé augmente, est ensuite caractérisé (section 3.5.1) avant de discuter de ses critiques (section 3.5.2).

Cette section se poursuit avec une discussion sur le non-équilibre. Car la thermodynamique classique – aussi appelée à juste titre la « thermostatique » – se limite aux systèmes à l'équilibre. La stratégie explicative générale de la thermodynamique du non-équilibre consiste à présenter les processus thermodynamiques comme un ensemble de flux et de forces similaires à ce qui est présent en mécanique des milieux continus. La détermination de ces flux et ces forces nécessite toutefois la résolution de plusieurs dizaines d'équations dont les paramètres sont généralement inconnus. Le concept d'entropie y joue néanmoins un rôle important en ce qu'il vient encadrer cette détermination. Puisque les phénomènes loin de l'équilibre, donc dans un état de non-équilibre, sont irréversibles et comportent des systèmes ouverts, la stratégie explicative présente les états stationnaires de ces systèmes loin de l'équilibre comme échangeant de l'entropie avec leur milieu de manière à atteindre une structure particulière, un certain niveau d'organisation. Le défi est donc de préciser cette stratégie en quantifiant entre autres ces échanges et en définissant clairement les niveaux d'organisation en jeu (microscopique, mésoscopique, macroscopique). Au demeurant, il s'agit d'une stratégie explicative potentiellement très puissante puisqu'elle prétend unifier des phénomènes extrêmement divers, des phénomènes de convection thermique aux origines du vivant. C'est en cela que le concept d'entropie rejoint celui d'émergence, car les phénomènes émergents seraient possibles *parce qu'ils* expulseraient leur entropie dans leur environnement. Aussi séduisant que cela puisse paraître, ces deux concepts sont ambigus dans ce contexte. (Bien entendu, les chapitres suivants approfondiront cette question.)

Mais il est entendu qu'aucune exception (empirique) n'a été trouvée au second principe de la thermodynamique dans sa formulation en termes de travail et de chaleur, de laquelle découle l'impossibilité d'un mouvement perpétuel de second type. Ainsi, à l'instar du principe de la conservation de l'énergie, « l'existence d'une loi aussi précise et aussi indépendante des détails d'un modèle particulier doit avoir un fondement logique indépendant du fait que la matière soit constituée de particules interagissant entre elles » (Lieb & Yngvason 2003 : 34). Mais il est possible d'exprimer ce principe sans faire appel au concept d'entropie, pour autant que les concepts de chaleur et de travail soient clairement établis. C'est pourquoi, encore une fois, le second principe de la thermodynamique sera distingué du principe entropique.

Le concept d'entropie est donc rempli de promesses mais reste à mains égard ambigu. D'où l'intérêt d'un examen critique. Il est en outre associé à des concepts d'intérêts philosophiques, comme l'irréversibilité, la causalité, la probabilité et l'accessibilité (observationnelle, causale, interventionniste). Il intervient au sein d'explications scientifiques particulières comme celles définissant les processus possibles à partir d'un état particulier, celles décrivant un processus vers un état d'équilibre, ou encore celles définissant la probabilité d'un état microscopique à partir d'un état macroscopique. Les phénomènes que tente d'expliquer le concept d'entropie et le second principe sont souvent complexes

et nécessitent des hypothèses simplificatrices parfois discutables. Les systèmes thermodynamiques, comme à peu près tout système physique d'ailleurs, peuvent difficilement être isolés de leur environnement, ou du moins pas de manière suffisante, et cela entraîne des difficultés autant dans leur représentation mathématique que dans l'heuristique des arguments visant à lui donner une définition claire. Le problème est d'autant plus pressant lorsque les systèmes ne sont pas à l'équilibre, ce qui semble être davantage la norme que l'exception. En même temps, les changements d'entropie peuvent renseigner sur l'interaction d'un système avec son environnement, car un système dont l'entropie diminue ne peut être isolé.

La discussion qui suit saura montrer que le concept d'entropie est fécond tant du point de vue scientifique que philosophique, en épistémologie en particulier, mais aussi pour la philosophie en général.

4.1 Irréversibilité

Le problème de l'irréversibilité est un problème relié au concept de symétrie : un phénomène irréversible montre une certaine asymétrie temporelle en ce que l'évolution vers le futur est différente de celle vers le passé, et cette différence est généralement caractérisée par rapport aux données de l'observation ou par rapport à une certaine théorie scientifique. La question se pose alors si une telle asymétrie vient des phénomènes se déroulant dans le temps ou plutôt du temps lui-même. Mais les propriétés attribuées au temps doivent bien être inférées et certains principes scientifiques sont parfois utilisés pour de telles inférences. C'est le cas notamment du second principe de la thermodynamique et de ce qu'on appelle la « flèche du temps » : la directionnalité du temps à la base de la différence intrinsèque entre le passé et le futur proviendrait de l'asymétrie thermodynamique du temps où les états futurs de l'univers présentent une entropie plus élevée que celle des états passés. Mais à peu près tout dans cette dernière phrase est sujet à caution³.

Cette asymétrie expliquerait le fait que nous puissions influencer le futur et non le passé et que nous ayons des souvenirs du passé mais non du futur. Une telle asymétrie, pourtant de sens commun et peut-être même de nature intuitive, peut aussi sembler déroutante si l'on considère qu'il n'y a pas une telle chose pour son analogue spatial. De plus, la mécanique newtonienne est exempte d'asymétrie temporelle puisque les processus qu'elle décrit et leur processus inverses sont tous deux en accord avec cette théorie. Malgré cela, il semble évident que le temps s'écoule dans une direction, que le passé est différent de l'avenir, que plusieurs processus ne peuvent être inversés. Ce n'est donc pas surprenant que le problème de l'irréversibilité et celui de la flèche du temps aient soulevé un grand nombre de débats tant philosophiques que scientifiques (voir Reichenbach 1956 ; Penrose & Percival 1962 ; Feynman 1965 ; Earman 1967 et 1986 ; Bunge 1968 ; Prigogine & Stengers 1979 ; Denbigh 1989 ; Zeh 1989 ; Mackey 1992 ; Hawking *et al.* 1993 ; Price 1996 ; Uffink 2001 ; Carroll & Chen 2004 ; Torretti 2007 ; Callender 2011 ; Weinert 2013).

L'asymétrie temporelle qui serait à l'origine de l'irréversibilité des phénomènes est aussi partie prenante des débats concernant la nature de la causalité, où la cause est censée toujours précéder l'effet. Partant, cette notion est centrale à certaines discussions à propos de la structure des explications

³ Dans les mots d'Earman (1974 : 15) : « *Of all the problems which lie on the borderline of philosophy and science, perhaps none has caused more spilled ink, more controversy, and more emotion than 'the problem of the direction of time.'* » Mais le doute n'est pas toujours au rendez-vous et cela est particulièrement évident chez Haddad, Chellaboina & Nersesov (2005 : 1-2) :

« *Since the system entropy increases, the entropy of a dynamical system tends to a maximum, and thus time, as determined by system entropy increase flows on in one direction only. Even though entropy is a physical property of matter that is not directly observable it permeates the whole of nature, regulating the arrow of time, and is responsible for the enfeeblement and eventual demise of the universe.* »

scientifiques en général, et en particulier de ce qu'il est convenu d'appeler l'asymétrie des explications. Par exemple, une cause expliquerait son effet, ou encore un événement passé expliquerait un événement ultérieur, mais pas l'inverse. C'est aussi le cas du second principe de la thermodynamique qui permettrait d'expliquer l'irréversibilité de certains phénomènes. Ainsi, à la question « pourquoi un phénomène thermodynamique comme le refroidissement d'une tasse de café s'est-il produit ? », la réponse immédiate est souvent « parce l'entropie d'un système comme la tasse doit augmenter (et une entropie plus élevée implique certaines caractéristiques observées dans le phénomène devant être expliqué, soit la tasse plus froide, et un environnement plus chaud) ». Mais cette irréversibilité qui semble si manifeste et dont la thermodynamique tente de rendre compte n'est apparente (justement) qu'au niveau macroscopique où les entités sont dites observables. Il n'est pas si évident qu'elle se manifeste au niveau microscopique et à plus forte raison si la seule description que reçoivent ses entités soit celle de la mécanique newtonienne qui est, rappelons-le, réversible (en un certain sens du moins). C'est pourquoi certains ont avancé que l'irréversibilité observée dans certains phénomènes n'était qu'une illusion, un sous-produit de nos limitations perceptives ou même cognitives, en bref un épiphénomène. Que les propriétés macroscopiques décrites par des variables continues et des lois irréversibles « émergent » du mouvement complexe d'un grand nombre de particules et décrit par des lois réversibles est une question fondamentale sur laquelle s'est acharné Boltzmann et qui demeure aujourd'hui, pour l'essentiel, ouverte⁴.

En effet, la difficile connexion entre les descriptions microscopiques et les descriptions macroscopiques est à l'origine de bien des tentatives novatrices de révision de ce qu'exige une telle connexion théorique, de ce qu'exige une réduction théorique selon le jargon. Il semblerait que les propriétés des entités à un certain niveau de descriptions ne se retrouvent pas à un autre niveau et qu'en conséquence leurs descriptions ne peuvent être considérées ni équivalentes ni assimilables les unes aux autres. Face à ce problème d'ordre plus métaphysique que méthodologique, on peut être tenté de revoir les différentes descriptions en jeu afin de faciliter cette connexion ou encore de définir un phénomène approprié faisant le « pont » entre celles-ci. À ce dernier égard, l'émergence est aux premières loges (Chapitre 2). À titre d'exemple, certains ont avancé que l'irréversibilité « émerge » comme une conséquence nécessaire des instabilités présentes dans les systèmes plus complexes (Coveney 1988) ; que pour les systèmes finis, une « hypothèse perpétuelle sur le désordre » est

⁴ Earman (1986 : 226) propose trois raisons pour le peu de progrès réalisé dans les travaux sur l'irréversibilité : (i) les chercheurs adoptent des approches avec différents présuppositions et prédilections ; (ii) les résultats mathématiques sont difficiles à atteindre ; (iii) le problème concerne autant l'explanandum que l'explanans, car il y a plusieurs définitions de l'irréversibilité. Il y a peu de raison de croire que l'état des progrès s'est amélioré depuis. Le livre édité par Hitchcock (2004) sur les débats contemporains en philosophie des sciences consacre d'ailleurs une section au « problème de la flèche du temps ». Pour ce qui est de la question de l'émergence, voir les Chapitres 2 et 6.

nécessaire pour avoir un comportement irréversible au niveau macroscopique (Wehrl 1991) ; que la réversibilité, devant l'irréversibilité observée, peut se manifester sur une échelle de temps « au-delà de l'éternité », c'est-à-dire de plusieurs ordres de grandeur supérieurs à l'âge de l'univers (Gallavotti 1999).

Cette section vise principalement à préparer la discussion sur l'irréversibilité telle que définie par le second principe de la thermodynamique et le principe entropique. Encore faut-il que ce principe soit clairement établi, ce qui sera l'objet de la section suivante. Mais il est impossible de discuter de la notion d'irréversibilité sans aborder celle du temps. Et il est évident que la question de la nature du temps ne pourra être traitée de manière exhaustive ici, ne serait-ce qu'en raison du volume considérable des discussions de nature métaphysique à son propos. Mais il est opportun de prendre position sur certaines questions philosophiques reliées au temps afin de prendre la pleine mesure du problème de l'asymétrie thermodynamique, indispensable à l'interprétation du concept d'entropie. D'abord, des distinctions s'imposent quant au sens et aux référents de ce concept d'irréversibilité.

4.1.1 Sens et référence

Le concept d'irréversibilité est utilisé avec plus ou moins de rigueur dans plusieurs disciplines scientifiques et dans le langage courant. Par exemple, la biologie l'emploie avec sa loi de Dollo où les organes « perdus » le sont à jamais, l'économie l'entend comme le degré auquel les effets d'une « mauvaise » décision peuvent être remédiés, la médecine comme l'impossibilité de guérir une maladie et l'éthique environnementale fait référence à des dommages « graves et irréversibles » dans son « principe de précaution » (voir Manson 2007). D'un point de vue plus formel, et plus près de son emploi en physique, le concept d'irréversibilité peut recevoir au moins deux interprétations (sens) ainsi que deux applications (références)⁵. Mais dans chacun des cas il s'agit bien d'une asymétrie par rapport au temps et non par rapport à l'espace. D'une part, l'irréversibilité peut se concevoir comme *non-invariance par renversement du temps* (non-IRT) ou comme *irrécupérabilité*, la première étant plus restrictive que la seconde car elle restreint davantage le domaine des possibles pour les états d'un processus. D'autre part, l'irréversibilité peut s'appliquer aux *processus* ou aux *théories*, comme quoi sa portée peut être ontologique ou épistémologique. (Cette discussion est inspirée de Uffink 2001 ; voir aussi Earman 1967 et 1974.)

⁵ Il est donc entendu que l'irréversibilité comme concept est une représentation mentale, qui a un sens et qui peut s'appliquer à des unités sémantiques, possédant des référents et une extension, où l'ensemble est susceptible de recevoir une valeur de vérité (voir Laurier 1980 et Panaccio 2011).

4.1.1.1 INVARIANCE PAR RENVERSEMENT DU TEMPS

Un *processus* est IRT si son « déroulement », c'est-à-dire l'ordre de la séquence de ses états, est similaire dans un sens temporel ou dans l'autre (par exemple si on le fait jouer sur film), de sorte que la séquence des états du processus sont les mêmes mais dans un sens inverse. Par exemple, le film d'une balle lancée en l'air montre des comportements similaires mais des trajectoires inversées si celui-ci est passé à l'endroit ou à l'envers, et, au surplus, dans chacun des cas on voit la balle accélérer avec la même accélération et dans la même direction, celle du sol (bien que leur vitesse soit inversée). Typiquement, un processus déterministe est IRT puisque son évolution peut être décrit par une bijection et son inverse par la réciproque de cette bijection⁶.

En revanche, une *théorie* est IRT si la classe des processus qu'elle permet est IRT. De manière formelle, ce critère peut être exprimé selon un espace des phases Γ contenant tous les états possibles d'un système, c'est-à-dire soumis aux contraintes de ladite théorie, où x est un point dans cet espace représentant un état instantané du système ; de sorte qu'un processus est représenté par une courbe paramétrée :

$$P = \{x_t \in \Gamma : t_i \leq t \leq t_f\}.$$

Il est alors possible de définir une classe W de tous les processus possibles selon la théorie en question, en d'autres mots l'ensemble des mondes possibles (selon la théorie), ainsi que la description d'un processus temporellement inversé P^* selon une transformation (involution) R inversant tous les signes des potentiels (moments et champs magnétiques), tel que

$$P^* = \{(Rx)_{-t} : -t_f \leq t \leq -t_i\}.$$

Il est alors acquis que les processus P et leur inverse P^* se déroulent dans le même temps, c'est-à-dire qu'ils ont la même durée. Ainsi, une théorie est dite IRT si la classe M des mondes possibles est close par rapport au renversement du temps, tel que

$$\text{Si } P \in M \text{ alors } P^* \in M.$$

À titre d'exemple, la mécanique classique est IRT, car si $\{x_1(t) \dots x_N(t)\}$, où x_i représente la position d'une des N particules, est une solution de ses équations dynamiques, alors $\{x_1(-t) \dots x_N(-t)\}$ l'est aussi. Ainsi, une théorie déterministe et IRT permet autant des prédictions que des rétrodictions, en d'autres mots elle présente une symétrie de prédictibilité.

⁶ Une bijection réciproque d'une bijection f est l'application qui associe à chaque élément de l'ensemble d'arrivée son unique antécédent par f (voir Mackey 1992). Il faut préciser toutefois qu'un certain processus ne peut pas toujours être inversé, en pratique, même si les lois de la physique n'interdisent pas un tel processus inversé.

Quelques commentaires. Ces critères s'appliquent aussi bien à une *loi* car une théorie pourrait en principe ne contenir qu'une seule loi. Par ailleurs, l'IRT d'une théorie ou d'une loi est aussi appelée *covariance* T . L'IRT d'une théorie est plus restrictive que sa contrepartie nomologique, c'est-à-dire par rapport à une loi, puisqu'une théorie peut bien entendu être un ensemble de lois. Ainsi, une loi peut être IRT, comme la seconde loi de la mécanique newtonienne, tandis que son application au sein d'une théorie ne le serait pas. Par exemple, une involution (R) de la mécanique newtonienne peut produire un processus temporellement inversé (P^*) par rapport à la trajectoire d'un corps quelconque où l'état final devient l'état initial et vice-versa. Cependant, si les différents potentiels, moments et champs magnétiques s'appliquant à ce corps, comme la gravitation, ne sont pas aussi inversés, alors il n'y aura pas nécessairement un processus temporellement inversé P^* présentant la séquences des états à l'envers d'un phénomène particulier. De plus, l'IRT d'une loi ou d'une théorie n'implique pas qu'il y ait invariance d'un avant et d'un après, c'est-à-dire des états initial et final, si l'on tient compte des conditions aux limites. Car ce n'est pas parce qu'une loi est IRT que l'état final va correspondre à l'état initial pour toute condition aux limites, incluant des processus pouvant s'interposer dans le processus décrit par la loi. Par exemple, une balle peut rouler d'un point a à un point b tandis qu'une inversion de la loi la fera passer de b à a ; mais si une barrière est placée sur la trajectoire avant le passage de la balle, l'inversion de a à b ne tiendra plus. Cette remarque est plutôt élémentaire, mais il faut aussi remarquer que des lois IRT n'élimine pas la possibilité d'assigner un ordre temporel, une asymétrie temporelle.

La référence de l'IRT, qui implique la possibilité d'une inversion de la séquence des états, doit donc être clairement précisée, mais ses implications doivent aussi être maîtrisées. Qu'une théorie soit IRT signifie qu'elle autorise certains processus et leurs processus inverses. Si la théorie est déterministe, alors elle offre une symétrie de prédictibilité. Mais cela n'implique pas que les processus le soient, car sa description doit être mise en conjonction avec des conditions initiales et le résultat peut évidemment être différent suite à une inversion de l'avant et de l'après du processus.

4.1.1.2 IRRÉCUPÉRABILITÉ

Un *processus* est dit *irrécupérable* si la transition d'un état initial x_i à un état final x_f obtenu lors d'un processus quelconque ne peut être complètement « défait » une fois le processus enclenché (par ex. le vieillissement). C'est-à-dire qu'il n'y a pas de processus commençant à x_f et se terminant à x_i . Il n'est donc pas possible de revenir à l'état initial et on dit que l'état ne peut être récupéré ou recouvré. On peut alors porter l'attention sur le processus reliant x_i à x_f ou sur l'état lui-même, leur attribuant ce caractère d'irrécupérabilité. Car si x_i est irrécupérable, on peut dire que le processus partant de cet état vers n'importe quel autre état est aussi irrécupérable. La question est autre de savoir si ce processus est responsable ou non de cette irrécupérabilité. Cette interprétation de l'irréversibilité serait due à Planck

(Uffink 2001 : 316 ; voir aussi Denbigh 1989). L'irréversibilité définie selon l'irrécupérabilité se distingue de la non-IRT sous trois aspects.

D'abord, pour qu'il y ait récupération en ce sens, il suffit de récupérer l'état initial x_i à partir de l'état final x_f sans égard à un quelconque processus P^* où chaque étape du processus original P serait retracée, ni à une loi réversible quelconque régissant le passage l'état initial à l'état final et vice-versa. L'évolution, selon la loi de Dollo qui stipule qu'un trait chez un ancêtre ayant disparu au cours de l'évolution ne réapparaîtra pas chez l'un de ses successeurs, est irréversible en ce sens. Comme il n'y a pas de caractérisation spécifique du processus, le critère de l'irréversibilité comme irrécupérabilité est un critère plus faible que celui non-IRT puisque seuls les états initial et final sont impliqués, et non pas la séquence complète des états du processus en question. En ce sens restreint de la récupérabilité, puisqu'on introduit généralement des conditions supplémentaires (dans ce qui suit), elle est impliquée par l'IRT.

Ensuite, l'état de l'environnement, soit l'ensemble des systèmes auxiliaires ou adjacents reliés au système à l'étude, doit aussi être intégralement récupéré ou recouvré. Un problème avec cette interprétation de l'irréversibilité est donc d'attribuer des caractéristiques à un système, définissant ainsi son état, en fonction d'autres systèmes auxiliaires dans l'environnement possédant des caractéristiques similaires. C'est-à-dire qu'il est « circulaire » (Uffink 2001 : 317) d'établir les conditions permettant d'attribuer des caractéristiques (par ex. de température, d'énergie et d'entropie) nécessaires pour définir l'état d'un système tout en présupposant que l'environnement les rencontrent déjà. Mais si l'on suppose qu'il est possible d'attribuer, dans la définition du système, un état \mathcal{Z} à l'environnement, on peut alors formaliser le critère de la réversibilité selon la récupérabilité en fonction d'un processus P s'il existe un processus P' tel que

$$\{x_f, \mathcal{Z}\} \rightarrow_{P'} \{x_i, \mathcal{Z}\}.$$

Il y a donc une condition supplémentaire à celle de la récupération de l'état initial. Cette condition peut être satisfaite en imposant que cette récupération soit effectuée sans intervention extérieure, sans que l'environnement ait un rôle à jouer. En bref, il faudrait que l'état initial soit recouvré à partir de l'état final et ce, *dans les limites du système*, c'est-à-dire que le processus inverse (P') soit contenu dans le même espace des phases. Cette condition est évidemment rencontrée dans le cas d'un système isolé.

Enfin, une troisième différence entre le critère de l'irréversibilité selon l'irrécupérabilité et la non-IRT concerne la notion de « possible ». Celle-ci reçoit au moins trois types de référents : « possible » peut renvoyer à la réalité, à une théorie ou encore aux capacités humaines. Il n'est pas suffisant que le processus P' soit permis par une certaine théorie, ou encore qu'il soit seulement à notre portée, il faut plutôt qu'il soit disponible, donc possible, dans le « monde réel ». Par exemple, s'il faut supposer que le

système est exempt de friction et qu'il n'est pas possible en pratique de l'éliminer, alors le processus n'est pas disponible et la condition n'est pas remplie. Un autre exemple plus éclairant dans la présente discussion est celui de la possibilité théorique d'inverser la vitesse et la direction de toutes les molécules d'un gaz, qui renvoie cependant à une impossibilité pratique, ce processus ne nous est pas disponible. Bien entendu, la signification de l'expression « monde réel » dépend d'un ensemble de considérations épistémologiques plaçant différentes exigences sur le processus en question. Par exemple, un tel critère pourrait exiger qu'un processus tel que P' soit observable de manière à rencontrer ce réquisit de réalité, autrement dit selon des critères faisant en sorte que « réel » signifie « épistémologiquement fondé par des moyens empiriques ».

Mais en général, l'irréversibilité est interprétée de manière moins stricte, de manière à ne pas imposer de condition aussi forte à l'environnement du système en question, généralement pour des raisons pragmatiques et par rapport justement à ces capacités humaines. Dans un contexte socio-économique par exemple, cette condition faisant référence à l'environnement peut se traduire par le « coût » nécessaire à la récupérabilité, à la « quantité d'intervention » demandée. Par exemple, un accident ou une maladie peuvent sembler irréversibles en ce sens où le coût nécessaire à la récupération de l'état initial (ou très près de l'état initial) est trop élevé, voire impossible dans les circonstances, de sorte que les moyens ou les ressources disponibles soient jugés insuffisants pour retrouver l'état de l'individu avant l'accident ou celui du patient avant la maladie. C'est en ce sens qu'on parle de rétablissement (« *recovery* »). Ce coût peut être interprété de bien des façons, selon, ici aussi, ce qui est possible en principe ou en pratique. Hormis cette notion de coût par rapport aux ressources déployées s'ajoute celle du temps, c'est-à-dire la durée nécessaire pour qu'un processus donné mène à l'état récupéré. Par exemple, la perte d'une quantité d'hydrocarbure comme le pétrole, par la combustion par exemple, peut sembler irréversible à l'échelle d'une vie ou de quelques générations, mais elle est récupérable sur plusieurs millions d'années car elle peut être produite par ce processus d'accumulation et de maturation de matière organique au sein d'une quantité de matière minérale. Un processus ou un état peut donc être jugé irréversible s'il est impossible de récupérer un état initial par les moyens disponibles dans l'environnement et le temps imparti pour y arriver.

Ainsi, l'irréversibilité selon l'irrécupérabilité n'implique pas et n'est pas impliquée par la non-IRT, de sorte que l'irrécupérabilité n'est ni une condition nécessaire ni une condition suffisante à la non-IRT. Il faut aussi préciser que la condition du retour à l'état initial de l'environnement \mathcal{Z} peut prendre une forme plus ou moins restrictive. Car à l'extrême, l'environnement \mathcal{Z} peut être l'univers au complet moins le système en question, en sorte que son retour à l'état initial semble impossible. Il semble plus raisonnable de limiter l'environnement \mathcal{Z} à l'ensemble des systèmes pertinents ou causalement reliés au système en question, celui à partir duquel le critère d'irrécupérabilité est évalué. Plus précisément, si un

système échange matière et énergie avec son environnement, la condition du retour à l'état initial de l'environnement devrait être évaluée en fonction de ces échanges. C'est-à-dire que les systèmes qui ne sont pas directement connectés au système, ceux qui n'ont pas servi à ce retour à l'état initial, ne sont pas considérés comme faisant partie de l'environnement considéré \mathcal{Z} . Encore plus spécifiquement, les variables théoriques déterminant un état particulier (par ex. un état thermodynamique) devrait être les mêmes que celles caractérisant l'environnement. De sorte que l'évaluation du respect du critère d'irréversibilité au sens d'irrécupérabilité dans le cas de la diffusion de la crème dans mon café n'a pas à tenir compte de la désintégration d'un atome de césium dans la galaxie d'Andromède, par exemple. Mais aussi et surtout elle ne tient pas compte des quantités de chaleur et de travail qui ne sont pas échangées entre le système et l'environnement, et tout autre variable thermodynamique qui ne sont pas influencées par ces échanges.

L'interprétation de l'irréversibilité au sens d'irrécupérabilité exige donc des précautions, car elle a plus à voir avec l'impossibilité de recouvrer l'état initial sous certaines conditions qu'avec une inversion d'un avant et d'un après ou d'un quelconque processus inversé pouvant les connecter. L'attention est focalisée sur la disponibilité d'un processus pouvant récupérer l'état initial sans trop modifier l'environnement. Les expressions « disponibilité d'un processus » et « sans trop modifier l'environnement » agissent ici à titre de paramètres. Dans le premier cas la disponibilité peut être déterminée par ce qui est considéré réel ou possible dans le monde qui est le nôtre, par une théorie ou par les moyens dont on dispose. Pour des raisons pragmatiques on peut imposer des contraintes temporelles et matérielles à la modification de l'environnement, comme quoi l'état devrait être récupéré avec des ressources et un temps donnés. Dans le second cas, l'environnement est plus raisonnablement défini non pas comme l'univers au complet hormis le système principal en question, mais comme le système auxiliaire étant modifié par des échanges avec le système principal et caractérisées par les mêmes variables théoriques que celles définissant l'état du système principal. Ainsi, supposant l'évolution temporelle d'un système à partir de conditions initiales qui seraient recouvrées après un certain temps, c'est l'idée de cycle qui se dégage ; mais l'idée d'un cycle irréversible est contradictoire à moins de tenir compte des changements dans l'environnement. De plus, il se peut qu'un état, pris comme une combinaison locale de l'ensemble de l'univers, soit reconstitué, mais cela n'implique pas, et il serait peut « probable » que ce soit le cas compte tenu du nombre de degrés de liberté en jeu, que l'état de l'environnement au grand complet, en bref de l'univers, le soit aussi. En somme, ce n'est pas l'événement qui ne peut pas revenir, c'est l'instant, entendu comme l'état de l'univers dans son ensemble (section 2.2.1).

4.1.2 Asymétrie temporelle

La question de l'asymétrie temporelle est bien entendu reliée à la question de la nature du temps, mais aussi de l'irréversibilité car un processus irréversible, et surtout pour des raisons de principe, renvoie à l'idée de la directionnalité du temps. Et bien entendu cette question occupe un large auditoire, de la physique des particules à la poésie en passant par la psychologie et la sociologie. Mais il est absolument vain ici de retracer l'ensemble des discussions à propos du temps, ni en se « limitant » au domaine philosophique (voir McTaggart 1908 ; Holmes 1960 ; Bunge 1968 ; Sklar 1993 ; Barreau 1999 ; Callender 2002, 2006 et 2011b ; Oaklander 2008), ni même en se concentrant sur les discussions touchant l'interprétation entropique de la direction du temps (pour des discussions plus détaillées, voir Reichenbach 1956 ; Earman 1986 et 2006 ; Price 1996 ; Albert 2000 ; Schulman 2010 ; Callender 2011a). Mais étant donné la pléthore de discussions identifiant le concept d'entropie comme l'ingrédient fondamental de l'asymétrie temporelle, il convient d'approfondir cette notion.

Il est aussi important de bien cerner le contexte de ce que prétend faire le concept d'entropie au sein de la thermodynamique et de son impact dans les discussions sur la nature du temps car elles influencent en retour l'interprétation de ce concept. On peut ainsi se poser la question de l'origine de l'asymétrie thermodynamique, comme quoi l'entropie ne peut qu'augmenter, mais on peut aussi prétendre que l'asymétrie thermodynamique explique les autres asymétries temporelles, comme celle de la causalité par exemple. L'asymétrie thermodynamique peut donc se présenter autant comme explanandum que comme explanans face au problème de l'asymétrie temporelle⁷. On peut en revanche prétendre que le temps lui-même est asymétrique et éviter ainsi de parler d'asymétrie dans le temps, bien que les termes de ce débat ne soient pas très clairs selon Sklar (1993). Mais peut-être que l'asymétrie thermodynamique n'est qu'une asymétrie *de facto* n'étant pas fondée dans une loi fondamentale asymétrique. Sa prétention d'expliquer d'autres asymétries fondamentales serait ainsi compromise. Il importe donc de bien préciser le point de départ et l'objectif visé car on explique généralement un phénomène ou un état de fait par un phénomène ou un état de fait jugé plus fondamental. Expliquer ainsi l'asymétrie causale par l'asymétrie thermodynamique place cette dernière en porte-à-faux de justifications épistémologiques importantes. Après tout, est-il possible de concevoir des phénomènes thermodynamiques sans la notion de temps ?

⁷ Évidemment, bien d'autres interprétations que celles s'appuyant sur le concept d'entropie ou le second principe de la thermodynamique ont été proposées et certaines d'entre elles seront présentées plus loin, d'autres seront laisser de côté. Mais on peut noter celles de Coe (1969) qui soutient que le temps est une « propriété générale de la matière » et un corollaire du *premier* principe de la thermodynamique.

4.1.2.1 LA QUESTION DU TEMPS

La question de la nature du temps est un sujet pour la métaphysique mais elle traverse aussi plusieurs débats contemporains en philosophie des sciences (voir, par ex., Callender 2002 ; Ridderbos 2002 ; Hitchcock 2004 ; Dieks 2006 ; Lyre 2008). Généralement, six principaux courants s’y s’affrontent, autour des questions suivantes : si le temps existe ou non, et si oui, quelle partie du temps existe entre le passé le présent et le futur, et si oui ou non le temps possède une existence indépendante des objets et des processus. Ainsi, on distingue d’abord l’*endurantisme* (associé à Héraclite), où tout change et où le temps est une suite continue de présent, les objets étant tridimensionnels avec une position *dans* le temps, et le *perdurantisme* (associé à Parménide), où la réalité se veut « une » et où les objets ont autant des parties temporelles que spatiales, ils seraient donc quadridimensionnels. Ensuite, le *présentisme* affirme que seul le présent existe, tandis que l’*éternalisme* soutient que toutes les parties temporelles (y compris le futur donc) existent. Enfin, le courant *relationniste* soutient que les caractéristiques du temps sont des constructions à partir des événements et de leurs relations ; la conception du temps est dans ce cas tributaire de l’idée que l’on peut se faire des objets et des processus. Le courant *absolutiste* en revanche soutient que le temps existe indépendamment des séquences d’événements qui prennent place en son sein ; le temps se conçoit alors comme un sorte d’arrière-plan sur lequel viennent s’appuyer les différentes descriptions du monde, comme celle fournie, par exemple, par la mécanique newtonienne. Le temps peut ainsi être caractérisé par lui-même ou par son contenu, soit les événements, selon plusieurs propriétés qui ensemble devraient englober, d’une part, les concepts de simultanéité, de succession et de durée, mais aussi, d’autre part, ceux de présent, de passé et d’avenir.

Selon l’emploi spécifique du temps associé à chacune de ces triades, les propriétés que l’on attribue au temps seront celles de réalité ou d’illusion (c’est-à-dire, le temps est-il réel ou non ?), de linéarité ou de circularité, de continuité ou de discontinuité, mais aussi d’hétérogénéité, d’asymétrie, d’anisotropie, etc. (discutées plus loin). Mais l’important ici est surtout d’identifier sur quoi s’appuient les inférences permettant l’attribution de telles propriétés. Car les discussions touchant la question de l’asymétrie temporelle s’articulent souvent autour de l’identification d’un élément dans la description des phénomènes, par le sens commun ou des lois scientifiques ou tout autre discours prétendument fondé, permettant d’inférer cette asymétrie. C’est le cas du second principe de la thermodynamique où la direction d’augmentation d’entropie des systèmes correspond à la direction du temps, du passé vers le futur.

La position adoptée quant à cette question est censée influencer et parfois même orienter la stratégie explicative globale des disciplines scientifiques en général, et en particulier de la thermodynamique. En effet, une définition du temps peut occuper une position très haute dans la hiérarchie d’une

axiomatisation théorique ou encore constituer le fondement d'une position ontologique, en sorte qu'elle peut offrir une sorte de cadre à la description des phénomènes. On considère souvent que tout phénomène est « soumis » au temps. Leur description serait ainsi soumise, logiquement, à la définition du temps. En retour, et comme mentionné précédemment, les théories scientifiques sont censées fournir des arguments à une définition particulière du temps. L'examen se concentre alors sur l'identification d'un élément de connaissance jugé suffisamment assuré pour caractériser indirectement à peu près toutes nos expériences, lesquelles sont censées justement s'inscrire « dans le temps ». Dans les deux cas, le résultat de cet examen est riche d'enseignement quant à la signification et à la portée des concepts scientifiques en jeu. Identifier la direction du temps du passé vers le futur à celle de l'augmentation de l'entropie, par exemple, révèle ainsi une prétention à l'universalité de ce concept, pour autant que le temps soit universel. En revanche, le réduire à nos impressions du passage du temps risque de confiner la direction du temps à une réalité subjective, voire à une illusion.

4.1.2.1.1 Caractérisation du temps

Plusieurs notions sont toutefois à distinguer quand vient la question du temps, car bien des discussions à son sujet n'ont produit que de la confusion. Il convient d'abord de distinguer les propriétés *intrinsèques* des propriétés *extrinsèques* du temps. Les premières réfèrent à l'entité même que serait le temps, tandis que les secondes renvoient aux phénomènes, à ce qui se produit dans le temps, c'est-à-dire aux changements opérés dans les phénomènes. La pertinence même d'une telle distinction peut être remise en cause si l'accès épistémique aux propriétés intrinsèques du temps est jugée impossible, si tout ce que l'on peut connaître du temps passe par les phénomènes. Mais il convient ici d'apporter d'autres précisions. Ainsi, les *événements* correspondent au « contenu » d'une position dans le temps, autrement dit une étape d'un phénomène ; une position dans le temps est un *instant*, et une suite d'instants définit une *durée* et correspond à un *moment*. Aussi, une série d'événements peut présenter un *ordre*, à l'instar des points sur une droite, mais cela n'implique pas nécessairement une *direction*, à l'instar de la suite des nombres entiers, présentant ainsi une *asymétrie*. La question se pose de savoir si la suite des instants est limitée ou non et si oui, dans quelle direction. On en revient presque inévitablement à la question de savoir si l'univers a un début ou une fin, et partant, s'il en va de même pour le temps.

Une fois ces distinctions faites, il convient de présenter le cadre d'analyse où se situe la caractérisation du temps, le distinguant de son homologue dimensionnel, l'espace. Car la question de l'homogénéité du temps n'est pas analogue à celle concernant l'espace puisque l'espace peut être homogène ou non relativement à ces trois dimensions, tandis que le temps est censé ne présenter qu'une seule dimension. D'ailleurs, comme la simultanéité est la négation de la succession, il est possible de « réduire l'ordre temporel à une suite unilinéaire » (Barreau 1999 : 1841) de manière à lui

imposer une direction. Si, de plus, l'on considère le temps comme continu, où les durées particulières peuvent se chevaucher à l'infini, alors il est possible de procéder à ce qu'on peut appeler une « spatialisation » du temps, comme quoi il aurait la même structure qu'une dimension de l'espace euclidien, de sorte qu'une durée puisse être mesurée par des événements réguliers et répétitifs. Ce procédé de mesure permet d'imposer des relations d'antériorité et de postériorité établissant une direction à l'ordre d'une série d'événements, et donc une *asymétrie*. Cette asymétrie est le propre d'une *succession*, où si A est antérieur à B , B ne peut être antérieur à A ; contrairement à une relation symétrique, comme celle d'adjacence, où si A est adjacent à B , B est aussi adjacent à A . Les relations d'antériorité et de postériorité qui forment une succession sont aussi considérées comme *transitives* en ce sens où, si un événement A est antérieur à un événement B , lui-même antérieur à un événement C , alors A est antérieur à C . Une notion de temps définie en ces termes – asymétrie et transitivité – implique ce qu'on peut appeler l'*anisotropie* du temps.

Si ce que révèle les lois ou encore l'observation ou peu importe ce qui peut prétendre à un discours épistémologiquement fondé, semble montrer une certaine préférence pour certains processus, alors il est commun d'attribuer au temps lui-même les propriétés d'anisotropie et donc d'asymétrie (souvent étiquetées par l'expression « flèche du temps », discutée plus loin). C'est pourquoi les discussions sur la nature du temps peuvent être aussi nombreuses que celles portant sur l'épistémologie des lois et théories scientifiques. Par exemple, le temps tel que défini par la théorie newtonienne est « absolu, vrai et mathématique », sans relation à rien d'extérieur. La théorie einsteinienne de la relativité restreinte a modifié cette conception puisqu'elle ne dissocie pas le temps comme dimension de sa contrepartie spatiale, car les mesures de temps et d'espace sont interdépendantes, pour l'incorporer dans ce qu'on appelle l'espace-temps, et fait de l'univers, une entité essentiellement intemporelle, quadridimensionnelle et « cinématiquement fixe » puisque le module du quadri-vecteur est invariant. Cet espace-temps est considéré comme une version de l'« univers-bloc », soit une représentation où tous les événements sont « là ensemble », au même titre que tous les objets de mon bureau ont leur position spatiale, puisqu'il n'y a pas de simultanéité globale, pas de présent universel (Dieks 2006). La situation est différente dans la théorie de la relativité générale où la géométrie de l'espace-temps est influencée (et non pas *déterminée* (Maudlin (2012)) par la distribution de l'énergie dans l'univers. En ce sens, la physique oscille entre une position absolutiste et relationniste du temps. Malgré ces exemples théoriques où la description théorique des phénomènes sert de base à la caractérisation du temps, il y a tout de même des restrictions à l'influence d'une telle description sur cette caractérisation : on ne saurait prétendre qu'une intensification (dans le temps ?) du changement (par ex. un processus se déroulant plus rapidement) correspond à une accélération du cours du temps. En outre, il importe de

souligner que ces théories peuvent recevoir plusieurs interprétations qui portent à conséquences sur une conception particulière du temps.

Cependant, la question de savoir si les propriétés d'asymétrie ou d'anisotropie s'appliquent bien au temps et non seulement aux processus peut, on le conçoit, faire naître un débat terminologique ou sémantique, à savoir quel est le référent du terme défini. Puisque caractériser une série d'événements ne revient pas nécessairement à définir le temps lui-même. On peut se demander en effet si la tentative de caractérisation du temps comme une entité indépendante des phénomènes ne se résume pas à une réification du temps, c'est-à-dire à transposer les propriétés d'une abstraction à une entité considérée comme réelle, où il y aurait confusion entre le temps et la description du changement, en quelque sorte le contenu du temps. Par un procédé semblant tenir à la fois de la synthèse et de l'hypothèse, on passerait d'une propriété universelle, ou plutôt jugée telle, des phénomènes à celle d'une « supra-entité » indépendante – le temps – censée encadrer ces mêmes phénomènes et désignée par un substantif. Le temps ne serait ainsi qu'un résumé du comportement plus ou moins nomique de certains phénomènes. En bref, il s'agirait ici de former une catégorie ontologique à partir d'une catégorie épistémologique. Ce genre d'inférence demande bien sûr une justification. Cependant, il n'est pas évident qu'une argumentation visant à définir le temps puisse faire autrement, car il est entendu qu'on peut difficilement former une catégorie ontologique sans faire appel d'aucune façon à nos outils épistémologiques, à moins de faire appel à des jugements *a priori* métaphysiques ou à un procédé similaire à la logique transcendantale kantienne où le temps serait une condition de possibilité à la connaissance. Cette astuce consistant à transposer les propriétés des phénomènes ou de l'entendement à celles du temps ne doit toutefois pas faire oublier d'où proviennent les inférences à ces propriétés. Voyons donc d'où sont tirées ces inférences.

Il y a deux principaux emplois attachés à la notion de temps qui découlent de ce qu'on pourrait appeler des « bases d'inférence », ou simplement des « justifications », c'est-à-dire des sources de connaissance, comme peuvent l'être dans leur acception la plus large la raison et l'expérience, ou encore les discours scientifiques et le sens commun. Car – j'insiste – le caractère général mais aussi fondamental du temps ne doit pas faire oublier l'argumentation visant à le définir, ni ce qui pourrait justifier l'attribution de certaines propriétés aux phénomènes en fonction de celles qu'on attribue au temps lui-même. Par exemple, d'aucuns pourraient soutenir que le temps est linéaire parce qu'une lettre arrive toujours après avoir été écrite, ou que celui-ci est cyclique parce que le moment fatidique de la page blanche revient inévitablement ; et inversement, que la direction du temps impose un « progrès » pour le meilleur et pour le pire, ou encore que si tout a été dit, tout est à redire car le temps est cyclique. Cela dit, en tout état de cause, tirer des arguments pour l'anisotropie du temps et conséquemment de sa directionnalité à partir de la description de certains phénomènes, de façon

nomique ou non, exige un bagage philosophique de taille, de manière générale, mais aussi, plus spécifiquement, l'universalité des phénomènes en question sans présupposer cette même directionnalité. Ce qui est en jeu en définitive est une caractérisation de la différence, si tel est le cas, entre le passé et le futur.

Le premier emploi est le temps vécu par la conscience, comme appréciation subjective – le « *tempus* ». Il est entendu, en ce sens, que la perception du « passage » ou du « flux » du temps est propre à chacun⁸. Ainsi, des expériences suggérant que certains événements se répètent peuvent être à la base de la croyance au retour éternel, tout comme d'autres peuvent pointer vers une finalité dans certains processus, soutenant plutôt la croyance en une direction du temps, voire en une fatalité. De fait, l'expérience nous dit que certains types de phénomènes se produisent plus souvent que d'autres, qu'ils semblent être favorisés par la nature ; par exemple, que les corps ont tendance à tomber, que les êtres vivants vieillissent et meurent, que le chaud va vers le froid, etc. L'expérience nous dit aussi que certains phénomènes ou événements semblent se répéter ; par exemple, que des dés lancés redonnent la même combinaison dans les mêmes proportions, que les saisons reviennent, que les élections semblent mener à la même corruption des élites, etc. Pour autant, le temps *per se* ne peut être appréhendé comme un phénomène brut. Comment ces expériences sont interprétées, jugées, évaluées, etc. peut faire l'objet d'un examen qui dépasse largement les ambitions de la présente discussion. Mais il faut noter qu'elles sont à l'origine d'intuitions extrêmement fortes qui servent souvent de base soit à la formulation d'une problématique particulière du temps (par ex. l'asymétrie de la mémoire et le problème de la directionnalité) soit à l'une de ses tentatives de solution. Ces intuitions sont reliées principalement à certaines asymétries qui peuvent se résumer à des cas particulier d'asymétrie temporelle, même si on peut tenter d'expliquer cette dernière par une autre asymétrie qui serait jugée plus fondamentale, et le problème consiste à savoir laquelle.

Parmi ces intuitions se trouve celle à la base de la définition de certains mots associés au temps et à la base même du langage avec ses « temps de verbe ». Le mot « événement », par exemple, trouve ainsi des synonymes avec « résultat », « fin », « dénouement », lesquels se laissent facilement appréhendés par quelque chose de passé, qui à la limite se décline au présent. En effet, il est tenu pour acquis que seuls les événements présents sont directement accessibles ; s'ils ne peuvent être que retenus, remémorés, regrettés, reconstruits, il faut les dire passés ; s'ils doivent être anticipés, attendus, planifiés ou prévus, il faut les dire futurs. Il semble donc y avoir déjà dans le langage quelque chose comme une direction du temps, déjà une asymétrie temporelle. La présence de cet ingrédient dans le langage

⁸ Par exemple, c'est ce que répond Rosalind à Orlando dans *As you like it* : « *Time travels in divers paces with divers persons.* » (Shakespeare 1994 : 293 [act. 3, sc. II])

n'apporte pas nécessairement une réponse à la question du temps, car il n'engage à aucun réalisme à son égard, mais il la rend particulièrement sensible aux équivocités et aux ambigüités. Sans doute que le langage est en mesure de rendre compte adéquatement de ces expériences décrites plus tôt, mais plusieurs intuitions ancrées dans le langage concourent à rendre vraies par définition certaines explications de l'asymétrie temporelle. Plusieurs notions associées au temps sont déjà asymétriques ; c'est le cas notamment avec celle de causalité, où une cause serait définie comme ce qui précède l'effet, et on pourrait facilement arguer dans le même sens avec des notions comme celles d'« influence », de « trace » ou de « souvenir ».

Il est donc commun de reconnaître différentes modalités d'existence aux événements passés, présents et futurs, ne serait-ce qu'en fonction de l'expérience subjective de la mémoire, de la conscience et de l'attente. Ces expériences perceptuelles, que l'on pourrait aussi nommer des états de conscience, constituent plutôt des bases non-inférentielles, c'est-à-dire qu'elles servent de prémisses qui ne sont pas démontrées parce qu'elles sont considérées comme immédiates. La question de savoir si elles méritent un statut épistémologique privilégié demeure cependant ouverte. Elles sont tout de même à l'origine de ces intuitions forgeant ces « morceaux de la sagesse populaire » qui peuvent être caractérisés par différentes asymétries temporelles. Ainsi peut-on noter les suivantes : l'asymétrie de la mémoire, car nous n'avons ni souvenir d'événements futurs ni d'expectative d'événements passés ; l'asymétrie de l'influence, car il semble que l'on puisse contrôler le futur et non le passé ; et l'asymétrie de la connaissance, comme quoi il est possible de connaître le passé mais pas l'avenir, du moins pas de manière aussi directe ou immédiate. Cette dernière est reliée de très près à l'asymétrie des « enregistrements » (« *records* »), comme des traces de pas ou des photographies, qui viennent semble-t-il toujours du passé mais pas du futur. Ces éléments de connaissance largement partagés militent en faveur de l'argument voulant qu'il soit, entre autres, irrationnel d'agir à l'encontre de ces asymétries ou de nier leur réelle effectivité. Or celles-ci sont toutes soumises à des oppositions de nature conceptuelle, épistémologique et métaphysique en ceci qu'elles portent sur la possibilité de la réduction d'un type d'entités ou de propriétés à un autre type d'entités ou de propriétés. On peut par exemple soutenir que les propriétés temporelles se réduisent aux propriétés causales, ou vice-versa.

Le second emploi a trait à l'ordre temporel et vise une description plus objective des phénomènes – le « *chronos* ». Tel que mentionné, la question de la nature du temps s'inspire de la structure théorique, concepts et lois, des différentes disciplines qui l'abordent. Par exemple, la physique relativiste montre qu'on ne peut dissocier le temps des moyens de le mesurer. C'est en effet la mesure, l'aspect quantitatif, qui joue ici un rôle important : des horloges suivant des trajectoires dans l'espace-temps différentes (donc avec des modules de quadri-vecteurs différents) mesurent des durées (intervalles de temps) différentes ; à l'instar des odomètres de voitures se déplaçant entre les deux mêmes locations mais selon

des chemins différents peuvent afficher des distances différentes (Maudlin 2012). Le temps peut ainsi correspondre à un paramètre d'ordonnancement définissant les relations d'antériorité et de postériorité, qui déterminent ainsi un ordre et une direction à une série d'événement. Cet emploi, plus opératoire mais aussi plus près de ce qui nous intéresse ici, sera discuté dans la prochaine section (voir aussi section 4.1.2.2).

4.1.2.1.2 Temps et historicité

C'est l'une des tâches anciennes de la philosophie de trouver un terrain d'entente, un « entrelacs » (Barreau 1999 : 1843), à ces deux emplois de la notion de temps. Bien que ce ne soit pas l'objet principal de la présente discussion, le premier emploi, celui du temps vécu, peut influencer de manière importante les discussions épistémologiques à propos de la signification et de l'interprétation des théories scientifiques, principalement par rapport aux asymétries temporelles, présentées plus tôt. Tel que mentionné, le second principe de la thermodynamique est ainsi souvent invoqué afin de rendre compte de l'asymétrie de la mémoire ou celle des « enregistrements » en général (voir Sklar 1993 ; Albert 2000 ; Norton 2000 ; discuté plus loin). Toutefois, même si ce principe parvenait à expliquer ces asymétries, il n'aurait pas nécessairement expliqué l'asymétrie du temps lui-même.

Car il ne va pas de soi que l'asymétrie du temps, définie, tel que discuté, par la direction des séries ordonnées d'événements, définissant à son tour les relations d'antériorité et de postériorité, puisse être déterminée strictement par l'observation des phénomènes, ni par les théories scientifiques. D'abord, cette asymétrie ou directionnalité est tout à fait compatible avec un monde régi par des lois IRT. En effet, dans un monde régi uniquement par des lois IRT à l'instar de la mécanique classique, il pourrait tout de même y avoir une « histoire » de ce monde (un peu plus dans un instant). Car la *possibilité théorique* d'inverser la séquence des états ne signifie pas qu'elle l'est ou le sera *dans les faits*. De sorte que l'observation ne rendra pas nécessairement compte de cette IRT. Par exemple, il est commun d'affirmer que toutes les lois fondamentales, celles sur lesquelles on peut raisonnablement se fier pour appuyer une métaphysique, seraient IRT, mais cela n'implique pas le retour, imminent ou non, des dinosaures. Plus simplement, ce n'est parce que le comportement des boules de billard est régi par une théorie IRT comme la mécanique classique qu'il nous est impossible de distinguer leur mouvement de gauche à droite ou inversement. En clair, l'IRT des lois n'implique pas l'IRT de la séquence des événements du monde. En conséquence, l'IRT des théories n'équivaut pas à une indifférence entre l'avant et l'après d'un processus.

Supposons alors un monde régi uniquement par des lois IRT à l'instar de la mécanique classique et qu'au surplus l'observation en témoigne peu importe la grandeur du système considéré. L'étude des phénomènes pourrait se résumer à des descriptions mécanistes similaires à celles concernant des boules

de billard en collision élastique. Le seul témoin d'un écoulement du temps serait notre mémoire ou un enregistrement quelconque de la succession des événements. Mais il n'y aurait rien dans les phénomènes qui laisserait présumer qu'un instant est fondamentalement différent d'un autre car les événements se reproduiraient, du moins on peut présumer qu'ils se reproduiraient plus souvent que dans notre monde, et cela serait d'autant plus vrai localement si l'on tient compte du nombre de processus contemporains et leurs interactions. Par exemple, une boule de billard peut suivre une trajectoire dont l'inverse est tout aussi possible en théorie, mais, ceci dit, il semble peu plausible que ce processus inverse (par ex. la boule se déplaçant de droite à gauche) soit aussi suivi d'un autre processus inverse connexe ou de toutes une série de processus inverses connexes (par ex. les aiguilles de l'horloge, qui se déplaçaient dans le sens ordinaire dans le cas du processus de la boule se déplaçant de gauche à droite, se déplaceraient alors dans le sens inverse, etc).

En revanche, dans le monde actuel, « notre » monde, des phénomènes comme le vieillissement ou un œuf se brisant au sol laissent présumer le contraire. Ainsi, dans un tel monde, les preuves empiriques appuyant la thèse d'une asymétrie temporelle serait plus minces si l'état du monde était réitéré, c'est-à-dire si l'état du monde, ou une large partie de celui-ci, était recouvert périodiquement et observé tel quel. Mais même dans un tel monde, la description des phénomènes ne serait pas nécessairement indifférente à l'inversion d'un avant et d'un après. Car le comportement d'un corps en mouvement comme une boule de billard peut être décrit par une loi IRT comme celles de la mécanique, mais la description peut être celle, par exemple, d'un mouvement allant de gauche à droite, montrant ainsi une asymétrie. Donc, si l'observation est d'une durée limitée, ce qui est *toujours* le cas, alors elle peut montrer une asymétrie suggérant une direction du temps, et ce même si la loi qui la régit est IRT. Il peut ainsi y avoir des raisons empiriques d'inférer une asymétrie temporelle malgré l'IRT des lois scientifiques. La question revient alors à identifier, si faire se peut, l'origine de cette asymétrie.

Pour que la direction du temps soit indétectable ou inobservable, au sens où la description des phénomènes, pour autant qu'elle soit adéquate, ne pourrait en rendre compte, et donc que rien dans les séries d'événements ne laisserait présager cette direction, au moins deux conditions sont nécessaires. Il faudrait, d'une part, que ces séries d'événements se produisent dans un sens comme dans l'autre, de sorte qu'il n'y aurait pas d'asymétrie dans les processus et il n'y aurait pas de phénomène qui semblerait être privilégié, comme le vieillissement par exemple. Rien ne laisserait présager une différence fondamentale entre la direction du passé vers le futur et la direction du futur vers le passé. Cette condition implique que les lois et les théories, que l'on suppose aussi soumises à la description des phénomènes, seraient IRT, tout comme les processus, à tout le moins ceux étant observés. Mais cette condition demeure insuffisante pour éliminer la possibilité de preuves ou d'indices empiriques à une asymétrie temporelle.

D'autre part, il faudrait que ces séries d'événements comprennent l'état de l'univers au grand complet, afin d'éliminer la possibilité d'avoir un « témoin », un « enregistrement » du changement, bref du cours du temps, où il y aurait au moins une série d'événements avec un ordre et une direction, au moins un processus dont la séquence des états ne serait jamais inversée. À titre d'exemple, un observateur regardant des processus IRT avec une horloge en mode normal aurait cette impression d'une direction du temps. En effet, cet observateur pourrait assigner à chaque événement une position temporelle dans une suite unilinéaire, à l'instar des points sur une droite (nonobstant la cardinalité choisie), correspondant à une heure précise, à un instant précis. Cette procédure pourrait même être appliquée en l'absence de ce « témoin » qu'est l'horloge, car un observateur doté d'une mémoire suffirait. L'observation des séries d'événements pourrait alors mener à la même procédure d'assignation d'une position temporelle où le recouvrement d'un état de l'univers le serait pour la première fois, puis la deuxième, etc. C'est en ce sens précis que j'entends le caractère d'« historicité » où tous les événements de l'univers comme un tout sont nécessairement différents, où l'asymétrie temporelle est manifeste. Il faut donc que la condition de l'état de l'univers au grand complet soit appliquée au pied de la lettre et que l'état initial de l'observateur soit lui aussi recouvert, de sorte qu'il n'ait même pas la mémoire des événements suivant cet état, autrement le recouvrement de l'état initial ne serait pas complet. La question qui s'impose alors est la suivante : l'historicité ainsi définie est-elle la direction du temps ?

Il va sans dire que ces conditions sont très difficiles à satisfaire. Leur non-satisfaction implique qu'aucun moment ou instant ne revient, à condition que la description des phénomènes observables soit adéquate. Mais si, au surplus, on pose l'hypothèse que leur satisfaction est impossible, alors il semble raisonnable d'associer caractère d'historicité à la direction du temps, car même si certains processus ou même certaines théories étaient réversibles, cette direction du temps serait garantie par l'impossible identité de deux instants de l'univers. En ce sens, on ne doit pas confondre le temps avec les phénomènes temporels, le contenant avec le contenu, la réversibilité locale et la réversibilité globale. Cela signifie aussi qu'on ne doit pas identifier le comportement d'une partie de l'univers avec celui de l'univers lui-même, auquel *le* temps s'appliquerait. Il va de soi qu'un monde sans changement, à l'instar de ce que représente une photo, ne fait aucun sens : ramener la diversité à une certaine unité n'élimine pas la possibilité ni l'effectivité du changement lui-même. La question est ailleurs.

Il faut plutôt dire pourquoi le temps *a* une direction, pourquoi *telle* direction, et que cela implique-t-il pour les phénomènes décrits par les théories scientifiques. Dès lors qu'il y a changement, soit une série d'événements inhomogènes, il est possible d'y assigner un ordre et une direction, donc d'y attribuer une transitivité et une asymétrie selon un paramètre d'ordonnancement. Cette procédure peut très bien correspondre à identifier le temps et son cours, bien que cette inférence demande une justification. Car

il ne faudrait pas confondre cette procédure à la description même des événements, laquelle peut présenter un changement régulier ou cyclique, ou encore un changement qui s'accélère. Elle peut ainsi mener à des conclusions contradictoires par rapport à l'inférence identifiant « quelque chose » – le temps – au terme de la procédure : à en juger par la description des phénomènes le temps serait tout à la fois à l'arrêt et en accélération, linéaire et cyclique, etc. En clair, la « typologie » des phénomènes ne correspond pas à celle du temps, ou ce qui est identifié comme le cours du temps à l'issue de cette procédure. Et entendons-nous bien : ce n'est pas parce cette procédure n'est pas effectivement menée qu'il n'y aurait pas de temps au sens ontologique du terme et encore moins que le changement serait nul. Mais même en ayant identifié la direction du temps à ce caractère d'historicité, qui se manifeste dans l'application de cette procédure, le raison fondamentale de cette direction demeure indéterminée, tout comme la question de savoir s'il pourrait en être autrement, si le temps pouvait être inversé.

Ainsi, puisque l'IRT des lois n'implique pas l'IRT de la séquence des événements du monde, il est toujours possible d'attribuer un ordre et une direction aux séries d'événements présentant du changement de manière irréversible, tel que discuté précédemment. En ce sens, une « histoire du monde », de l'univers, demeure possible. C'est-à-dire qu'il y a au moins une série d'événements avec un ordre et une direction, donc présentant une asymétrie. En conséquence, distinguant l'état de l'univers de celui d'une de ses parties, la description d'un phénomène se produisant dans l'une de ses parties est toujours possible, en opposition à celle d'une autre partie et même de l'univers au complet, ou presque, de sorte que cette description peut toujours être située par rapport à une histoire du monde et donc à cette procédure attribuant un ordre et une direction aux séries d'événements. C'est pourquoi on peut bien dire que le temps s'écoule même s'il ne se produit aucun changement dans une certaine partie de l'univers. Autrement dit, le « cours du temps » est différent de la « flèche du temps » (section 2.2.3). Mais que cette historicité soit présente n'implique pas qu'il n'aurait pu en être autrement, que l'ordre des choses n'aurait pu être inversé, au sens où un principe ou une loi est IRT, mais seulement qu'il en est ainsi. Dans le même exemple de la boule de billard, celle-ci peut se déplacer de gauche à droite sans que cela n'implique que la loi qui régit son comportement interdise un autre mouvement, à moins bien sûr de prendre en compte les conditions initiales de ce mouvement. Or, c'est justement ce qui est fait lorsque les conditions initiales sont appliquées à la description de l'origine de l'univers et à partir de laquelle son évolution générale peut être prédite. L'historicité du monde, et partant la directionnalité du temps, deviennent alors patentes, mais cela est fait à partir de cette description faisant intervenir des lois IRT et des conditions initiales, et non pas seulement à partir de ces lois qui ne présentent aucune asymétrie. Enfin, il faut mentionner cette possibilité, qui n'a pas été écartée, celle à l'effet que les changements observés et enregistrés lors de la procédure d'assignation d'un ordre et d'une direction

temporelle ne seraient que des illusions et qu'en fait, sous les apparences, il n'y aurait qu'une seule réalité permanente.

En conclusion. Il ne faut pas confondre l'asymétrie des processus avec celle du temps, ce qui revient à assimiler les propriétés extrinsèques du temps à ses propriétés intrinsèques. Mais, attendu que ce qui est identifié comme le temps n'est pas directement accessible, il est courant d'hypostasier en quelque sorte les propriétés des processus en une entité qualifiée sous le vocable « temps ». La légitimité d'une telle « hypostase » ou réification est un sujet pour la métaphysique. Il importe donc de justifier l'attribution de propriétés au temps à partir des propriétés des processus, si tant est que l'observation et les théories scientifiques constituent des bases d'inférence à de telles attributions. La même exigence vaut pour l'inférence quant à l'existence ou non du temps. Pourtant, le temps, au sens ontologique du terme, est censé fournir un cadre ou une forme aux phénomènes, et non l'inverse. Si le temps est posé comme premier et si par suite les phénomènes doivent refléter ses caractéristiques, il ne reste qu'un pas à franchir, un « pas kantien » pourrait-on dire, avant d'associer la catégorie du temps à une condition de possibilité de la connaissance de ces mêmes phénomènes.

Nonobstant cette possibilité, tant qu'un enregistrement demeure possible, comme la mémoire par exemple, une « histoire du monde », comme la description d'une suite d'événements et donc d'instantanés différents, est possible. Il est alors possible de procéder à l'attribution d'un ordre et d'une direction aux séries d'événements présentant un changement. Attribuer à ces changements dans les phénomènes une raison plus fondamentale, une cause commune, une loi ou un principe, peut alors revenir à définir la directionnalité du temps. Mais il semble qu'on ne soit jamais en mesure de justifier pleinement cette raison fondamentale. Par exemple, identifier la causalité comme la responsable de cette directionnalité semble éluder la question si tant qu'il s'agit déjà d'une notion temporelle, tandis que le second principe de la thermodynamique tente d'établir une portée universelle à un concept provenant d'une théorie qui ne l'est pas. En supposant que le second principe ou le principe entropique aient bien l'universalité nécessaire, deux obstacles demeurent avant qu'ils puissent jouer le rôle d'explanans de l'asymétrie temporelle. D'abord, il est possible que l'asymétrie thermodynamique soit simplement *de facto*, une espèce de sous-produit d'autres lois plus fondamentales ou même une conséquence des propriétés intrinsèques du temps lui-même. Ensuite, puisqu'un monde régi par des lois IRT est compatible avec son caractère d'historicité, on peut se demander si l'identification d'une loi non-IRT permet vraiment d'expliquer l'asymétrie temporelle.

4.1.2.2 LE TEMPS EN PHYSIQUE

Face aux difficultés d'expliquer l'asymétrie du temps et donc de justifier le rôle d'explanans du principe entropique, il convient d'analyser le rôle que le temps peut jouer en physique. D'abord, il a été mentionné qu'il ne fallait pas confondre les propriétés du temps avec celles des processus. Ensuite, l'idée très générale que l'asymétrie temporelle procède de l'asymétrie des processus physiques est difficile à déloger, bien qu'elle n'ait pas encore reçue de démonstration convaincante. Car il ne va pas de soi que l'on puisse identifier le temps au changement en général, ni sa mesure au changement régulier plus particulièrement. Les développements de la physique moderne ont compliqué les questions et les réponses par rapport au temps. En simplifiant un peu, on peut dire que la physique est passée d'une conception absolutiste, celle de Newton, à une conception relationniste, celle d'Einstein (pour plus de nuances, voir Earman 1970 et 2002 ; Dieks 2006 ; Maudlin 2012). Mais l'idée qu'il faille, face à la problématique du temps, interroger les phénomènes tels que décrits par nos meilleures théories scientifiques, était déjà présente avant les développements de la physique relativiste avec la problématique de la réversibilité des lois physiques et de l'apparente irréversibilité des phénomènes. À cet égard, le second principe de la thermodynamique prenait, et prend encore, la forme d'un singulier protagoniste.

Il a été vu précédemment que chaque théorie physique pouvait donner au temps un statut original et particulier. Alors que la théorie newtonienne fait du temps une entité indépendante et absolue, la théorie de la relativité restreinte et celle de la relativité générale modifient cette image à plusieurs points de vue. En effet, la cinématique de la relativité restreinte, qui repose sur des transformations mathématiques permettant de traduire l'état d'un système d'un référentiel à un autre, « imbrique » en quelque sorte le temps et l'espace, qui dépendent alors de la vitesse du système, car plus celui-ci va vite, plus son temps se « dilate » et son espace se « contracte » : plus précisément, la transposition du résultat d'une mesure d'une distance ou d'une durée dans un autre référentiel exige une transformation, dite de Lorentz, impliquant un facteur qui dépend de la vitesse relative des référentiels. Non seulement le temps perd son idéalité absolutiste, mais la simultanéité et donc la mesure du temps deviennent relatives au mouvement des observateurs. Dans la relativité générale, la gravitation est associée à une propriété géométrique de l'univers dont la structure spatio-temporelle est modifiée localement par la densité d'énergie des systèmes qui s'y trouvent. En outre, la cosmologie nous informe que l'univers aurait une histoire dont les origines se confondent avec les limites de la physique elle-même. Enfin, en mécanique quantique, les lois réversibles décrivant la dynamique des systèmes sont incapables de rendre compte du phénomène irréversible de la mesure. Il serait donc mal avisé de prétendre avoir en main *le* temps de la physique.

Mais, en physique comme en général, il y a deux principaux emplois de la notion de temps, quoiqu'il ne soit pas question ici du temps vécu : le *cours du temps*, ou simplement le temps, et la *flèche du temps*, appelée parfois aussi le devenir⁹ (voir Bunge 1968 ; Friedman 1983 ; Klein 1998, 2004 et 2010 ; Isham & Savvidou 2002).

Le premier emploi, déjà mentionné, est celui d'un paramètre d'ordonnement des états selon une logique temporelle, où l'on peut attribuer une valeur de vérité à des propositions en fonction de ce paramètre, comme dans la proposition « P à t_1 puis Q à t_2 » (où $t_2 > t_1$). Il s'agit alors d'une grandeur quantifiable susceptible d'ordonner des expériences et de les relier mathématiquement en leur attribuant une position numérique sur une « ligne du temps », donc en leur attribuant un moment. C'est pourquoi on parle parfois de la « spatialisation » du temps (déjà discuté), en ce sens où il représente une dimension d'un espace des phases. De ce point de vue, le temps peut être considéré comme un « producteur de durée » où le cours du temps a un ordre et une direction, même si celle-ci est arbitraire. La conséquence de cet ordonnancement est que chaque événement est relié aux autres par l'une des relations d'antériorité, de postérité ou de simultanéité. De plus, pour chaque moment donné, pour chacune des positions temporelles auxquelles sont assignés certains événements, deux ensembles d'événements distincts, l'un représentant le passé et l'autre le futur, sont définis. Cette « production de durée », que l'on désigne aussi métaphoriquement par l'« écoulement du temps », peut être régulier ou non, intrinsèquement ou non. Tel que mentionné, la mécanique classique propose un cadre rigide où cet écoulement est régulier, tandis que la relativité restreinte stipule que le temps, dans son acception objective quant à sa mesure empirique, dépend du mouvement de l'observateur procédant à cette mesure, l'écoulement étant alors irrégulier, ou « élastique » (Luminet 1994 : 60). Cet emploi est donc compatible tout autant avec une position absolutiste que relationniste.

Le second emploi est celui d'un paramètre d'évolution qui se présente dans la description de la dynamique des systèmes. Il s'agit d'une mesure de « comment les choses changent », une mesure du « devenir ». Qu'est-ce que cela signifie ? Que plusieurs lois scientifiques sont exprimées en fonction d'une variable indépendante par rapport à laquelle les changements d'état sont déterminés ; par exemple, la position \mathbf{r} d'une particule est déterminée par $\mathbf{r}(t) = \mathbf{r}_0 + \mathbf{v}t$. Il ne s'agit donc pas d'une mesure du changement à proprement parler, car l'état peut être recouvré suite à l'application d'une certaine valeur de la variable t , après « un certain temps », comme dans le cas d'un oscillateur harmonique, par exemple. Ce paramètre d'évolution décrit plutôt l'évolution du système, en ceci qu'il détermine, par la loi à laquelle il est appliqué, la séquence des états du système. Mais cette séquence

⁹ Isham & Savvidou (2002) parlent plutôt du « temps de l'être » (« *time of being* ») et du « temps du devenir » (« *time of becoming* »).

peut aussi être présentée en faisant abstraction de ce paramètre, en insistant sur le changement qu'a subi le système, lequel peut être plus ou moins prononcé, de sorte que l'intensité de ce changement soit prise pour du cours du temps lui-même ou son rythme. Mais est-ce que le temps s'accélère si l'on fait les choses plus rapidement ? Clairement, cette astuce impose des ajouts conceptuels inutilement. Cependant, il demeure conceptuellement possible d'attribuer au temps non pas une mais deux dimensions, dont l'une pourrait être soumise à un écoulement irrégulier, une accélération locale selon les phénomènes.

Avec ces deux emplois de la notion de temps, la physique ne rejette donc pas la possibilité d'identifier un cours du temps, et ainsi une histoire du monde et ce, bien que la plupart de ses lois soient IRT. Si une loi ou un ensemble de lois IRT, comme c'est le cas avec la mécanique classique, déterminent une série d'événements, en bref un processus, alors la série d'événements dans la direction inverse est tout aussi compatible, nonobstant les conditions aux limites, avec cette loi ou cet ensemble de lois. La seule façon de discriminer ces deux processus quant à une direction du temps qui serait considérée comme la « bonne » est de se référer à un « témoin », un « enregistrement », bref un processus irréversible ou simplement « dirigé », ou considéré tel. Il semble tout à fait plausible de caractériser, *prima facie*, un « enregistrement » comme ce qui est relié à l'événement original par un certain mode de détermination¹⁰, la causalité en étant un exemple. Puisque l'application d'une mode de détermination est temporellement asymétrique, soit intrinsèquement, comme la causalité où la cause précède toujours l'effet, soit par la conjonction de conditions initiales (déjà discuté), il est commun de qualifier un « enregistrement » comme étant, partiellement du moins, irréversible (minimalement au sens de non-IRT). La mémoire est un tel exemple. Mais il est aussi possible d'établir une direction du temps à partir de processus issus de lois IRT si la condition de la « comparabilité locale de l'ordre du temps » (Reichenbach 1956 : 35) est respectée. Cette condition stipule qu'il est possible d'établir empiriquement la direction du temps¹¹ de deux processus en juxtaposition spatiotemporelle. Cette possibilité découle de la détermination par les lois IRT des événements constituant le processus selon un ordre précis. Par exemple, deux boules de billard entrant en collision sont donc en juxtaposition spatiotemporelle : la description physique du processus global, incluant les deux trajectoires, n'aurait aucun sens si seulement l'une des deux trajectoires devaient être inversée et pas l'autre ; autrement dit, si l'une d'entre elles est fixée, l'autre l'est aussi.

¹⁰ Bunge (1959) définit un « déterminisme générique » comprenant plusieurs modes de détermination : « *All that is needed in order to maintain determinism in a general sense is to hold the hypothesis that events happen in one or more definite (determinate) ways, that such ways of becoming are not arbitrary but lawful, and that the process whereby every object acquires its characteristics develop out of preexisting conditions.* » (13). Ainsi, rien ne se produit de rien ni ne se produit de manière inconditionnelle et irrégulière, anomique et arbitraire.

¹¹ Reichenbach (1956) ne semble pas faire de distinction entre *ordre* et *direction* du temps.

Toutefois, ces mêmes lois, au premier chef celles de la théorie de la relativité, imposent des contraintes à la procédure d'assignation d'une position temporelle (déjà discutée). L'une des conséquences les plus remarquables de cette théorie concerne le caractère relativiste du temps découlant de sa mesure, car la description d'un processus dépend de la relation cinématique entre ce processus et le référentiel à partir duquel est établie cette description. On peut dire alors que la « cinématique est indissociable de la dynamique » (Paty 1994 : 55). Ainsi, la relativité de la simultanéité exige des corrections mathématiques aux signaux entre deux points de l'espace-temps, autrement dit il faut appliquer des transformations de coordonnées aux lois physiques entre plusieurs référentiels, qui sont des systèmes de coordonnées. Mais la théorie de la relativité préserve ce que d'aucuns appellent le principe de causalité et d'autres le principe d'antécédence, comme quoi tout mode de détermination (causal, stochastique, etc.) va du passé vers le futur. D'ailleurs, malgré l'interprétation de l'univers-bloc, où le temps semble être inexistant ou encore où tous les instants sont aussi réelles (éternalisme), il est commun de parler de la « structure causale de l'espace-temps ». De plus, l'« imbrication » de l'espace et du temps au sein de l'espace-temps ne prive pas le temps de ses privilèges, car la théorie de la relativité n'élimine ni la distinction conceptuelle du temps et de l'espace, ni le concept de signal électromagnétique, qui demeure un processus (Bunge 1967b : 206). En ce sens, elle ne remet pas en cause la directionnalité du temps et ne lui apporte pas non plus d'explication. (Pour plus de détails sur l'impact de la relativité sur la notion de temps, voir Earman 1970 et 2002 ; Friedman 1983 ; Paty 1994 ; Dieks 2006 ; Petkov 2007 ; Heller 2009.)

En conclusion, du point de vue de la physique, la réversibilité des lois signifie que la description nomique d'un processus dans un sens n'est pas plus « correcte », pas plus « physique », que la description dans le sens inverse. Mais cela ne signifie pas qu'elle élimine la possibilité que le temps ait un cours et donc une direction. Cette direction est présente au sein des deux emplois de la notion du temps en physique comme paramètre d'ordonnancement ou comme paramètre d'évolution, lesquels ne sont pas nécessairement incompatibles. La question n'est pas tant de savoir s'ils peuvent être assimilés l'un à l'autre ou si des propriétés particulières du temps comme la circularité les autorisent ou non. Plutôt, si la description du changement, et donc la physique elle-même, est impossible sans l'inscrire dans ce que l'on désigne avec ces deux paramètres, alors le prix à payer est trop élevé : ce serait alors notre conception du temps qui serait à revoir plutôt que de rejeter la physique.

4.1.2.3 FLÈCHE DU TEMPS

Bien que le temps ne soit pas directement perceptible, le devenir comme changement dans les processus peut l'être. Eddington (1928 : 69) a introduit l'expression « flèche du temps » (« *time's arrow* ») pour décrire la *directionnalité des processus* :

Je vais utiliser le terme « flèche du temps » pour exprimer cette propriété de sens unique du temps qui n'a pas son pareil dans l'espace. Il s'agit d'une propriété particulièrement intéressante d'un point de vue philosophique. On doit remarquer que —

- (1) Elle est vivement reconnue par la conscience.
- (2) Elle est également appuyée par notre jugement, qui nous dit qu'une inversion de la flèche rendrait le monde extérieur incompréhensible.
- (3) Elle n'apparaît pas dans les sciences physiques, excepté dans l'étude de l'organisation d'un certain nombre d'éléments. La flèche indique alors la direction d'un accroissement progressif du caractère aléatoire.¹²

Eddington fait ainsi référence à des inférences sur les modes phénoménologique, épistémologique et théorique pour justifier cette notion de directionnalité, laquelle mérite examen.

D'abord, lorsqu'il parle de la conscience, sans doute fait-il allusion à la phénoménologie de la durée, au temps vécu discuté précédemment, où la mémoire permet d'établir des relations d'antécédence et de postériorité aux événements perçus. En ce sens particulier, il y aurait bien une direction au temps, et donc une asymétrie, établie par l'ordre des séries de ces événements. Ensuite, lorsqu'il parle de notre jugement « qui nous dit qu'une inversion de la flèche rendrait le monde extérieur incompréhensible », sans doute a-t-il en tête une conception kantienne du temps comme une représentation nécessaire servant de fondement à toutes les intuitions, soit une « condition de l'existence des choses comme phénomènes » (Kant 1787). Mais il est étrange qu'Eddington parle d'une inversion de la direction du temps plutôt que sa disparition, car un ordre inversé offrirait tout de même un fondement à nos intuitions comme semble l'exiger la conception kantienne. Seulement, ces intuitions procèderaient à l'envers, mais il y aurait tout de même une direction du temps. Enfin, la dernière remarque sur l'« accroissement progressif du caractère aléatoire » est plus difficile à cerner. Il semble bien qu'il fasse référence aux systèmes macroscopiques composés d'un grand nombre de particules dont les trajectoires sont impossibles à calculer avec la mécanique classique, ce qui exige le recours aux statistiques. La thermodynamique semble bien établir cette « flèche » avec le second principe et il s'agit probablement de l'exception physique à laquelle Eddington fait allusion. La direction du temps ne serait ainsi

¹² I shall use the phrase "time's arrow" to express this one-way property of time which has no analogue in space. It is a singularly interesting property from a philosophical standpoint. We must note that —

- (1) It is vividly recognised by consciousness.
- (2) It is equally insisted on by our reasoning faculty, which tells us that a reversal of the arrow would render the external world nonsensical.
- (3) It makes no appearance in physical science except in the study of organisation of a number of individuals. Here the arrow indicates the direction of progressive increase of the random element.

apparente qu'au niveau macroscopique où les systèmes sont composés « d'un certain nombre d'éléments ». On constate donc qu'il y a plusieurs notions entremêlées, qu'il peut facilement y avoir confusion, et que cela n'est pas étranger aux nombreuses discussions sur le sujet.

Eddington fait mention d'au moins deux sources à l'expression de la « flèche du temps ». La première étant la « flèche psychologique », qui est celle de notre perception du « passage » du temps, comme dans l'exemple de nos souvenirs des événements passés contrairement aux événements futurs. La seconde étant la « flèche thermodynamique » où il y a accroissement de l'entropie et (donc ?) du caractère aléatoire ou du désordre. À cela on pourrait ajouter la « flèche cosmologique » où il y a dilatation de l'univers. Alors que signifie vraiment l'expression de la « flèche du temps » ? Il a déjà été mentionné qu'elle référerait à la directionnalité des processus, c'est-à-dire à l'asymétrie dans la séquence des états définie par un ordre et une direction. En effet, j'ai soutenu plus tôt que l'asymétrie dans le comportement de certains systèmes ne devait pas être confondue avec celle du temps lui-même. Nonobstant la possibilité de soutenir une thèse sur la circularité du temps à partir d'aprioris métaphysiques, j'ai aussi montré qu'une histoire du monde associée au cours du temps demeurerait possible malgré l'IRT des lois et malgré l'asymétrie de certains processus. Il faut donc, comme le fait la physique, distinguer le cours du temps de la flèche du temps (plutôt : la « flèche des processus »).

Toutes les asymétries temporelles rencontrées jusqu'à maintenant ont reçu des tentatives d'explications par d'autres asymétries jugées plus primitives ou plus fondamentales. C'est le cas notamment de la grande asymétrie du temps lui-même comme entité propre qui gouvernerait toutes les autres asymétries, « la flèche-mère de toutes ces fléchettes » (Klein 1998 : 5). Car il va de soi, semble-t-il, qu'une asymétrie ne peut être expliquée que par une autre asymétrie (van Fraassen 1989 : 233). Il semble donc y avoir une différence de nature ou à tout le moins de registre entre symétrie et asymétrie, en ce sens où il est permis de passer de l'un à l'autre que par un méta-principe encadrant ce passage¹³. Le problème se résume ainsi à savoir où et comment introduire cette asymétrie considérée comme primitive. Notons qu'il semble aller aussi de soi qu'une asymétrie demande une explication, mais que ce ne soit pas toujours le cas d'une symétrie, sans doute en raison de sa proche relation avec les notions d'ordre, d'harmonie et de beauté¹⁴. Plusieurs asymétries temporelles, plusieurs « flèches du temps », tirées de la physique moderne sont avancées comme candidates, soit comme explanans soit comme

¹³ De manière analogue dans un tout autre débat, une norme ne peut être justifiée que par une autre norme, à moins d'avoir en main un méta-principe stipulant la validité d'une norme ou autorisant le passage du descriptif au normatif (bref, en autorisant l'erreur naturaliste).

¹⁴ Ces raisons sont peut-être discutables, mais je souhaite simplement signaler que nous avons davantage tendance à accepter une symétrie comme « fait brut » (ne nécessitant pas d'explication) qu'une asymétrie. Par exemple, la science grecque considérait le mouvement circulaire comme étant le mouvement géométrique parfait, « donc » naturel, « donc » ne nécessitant aucune explication particulière. (Voir aussi Stewart 2007.)

explanandum ; par exemple : la flèche radiative (on n'observe qu'un seul type d'onde parmi les deux que décrit la théorie électromagnétique), la flèche quantique (entre évolution mécaniste et mesure), la flèche cosmologique (expansion de l'univers), et bien sûr la flèche thermodynamique (voir Luminet 1994 ; Price 1996 ; Lyre 2008 ; Callender 2011a).

Il a été vu que l'asymétrie thermodynamique était considérée soit comme un explanandum de l'asymétrie temporelle soit comme son explanans. L'asymétrie thermodynamique est définie à partir du second principe exprimé selon le concept d'entropie, soit le principe entropique, qui stipule en termes simples que l'entropie ne peut pas décroître : lors d'un processus passant d'un état d'équilibre A à un autre état d'équilibre B , la différence d'entropie de ces deux états, $S(A) - S(B)$, est supérieure ou égale à une certaine quantité définie positive par convention. Puisque la thermodynamique est une science phénoménologique (se limitant à l'observable), l'entropie thermodynamique est tributaire en grande partie des constatations empiriques qui sont, au demeurant, plutôt solides en ce qui la concerne (section 3.5). Néanmoins, on peut bien affirmer qu'il n'y a pas de lois fondamentales, ou du moins jugées *plus* fondamentales que le second principe, qui interdisent de comportement allant à l'encontre de l'asymétrie thermodynamique, ou même qu'en l'état le second principe n'inclut aucune asymétrie temporelle puisqu'il ne fait pas référence aux états de non-équilibre. Il faut bien l'admettre, si un système passe d'un état à un autre, qu'ils soient à l'équilibre ou non, nécessairement il fut en dehors de l'équilibre à un moment ou à un autre durant ce passage d'un état à un autre état, car la notion de processus quasi-statique demeure *stricto sensu* une idéalisation. Malgré plusieurs objections auxquelles peut faire face le principe entropique (section 4.2.5), il n'en reste pas moins qu'il n'est pas d'emblée écarté comme un principe faux ou falsifié. Plutôt, ces objections visent à revoir son statut épistémologique, en le distinguant par exemple d'une loi au sens stricte, ou en révisant sa portée, comme quoi il ne s'appliquerait pas aux processus pertinents par rapport au problème de l'asymétrie temporelle (discuté plus loin).

Le besoin d'expliquer l'asymétrie thermodynamique, qui est un exemple de l'irréversibilité *qua* non-IRT de la théorie, est particulièrement manifeste lorsque celle-ci est opposée à la symétrie mécanique, à la réversibilité des lois mécaniques. En effet, si l'on admet qu'un comportement global, manifesté à notre échelle, au niveau macroscopique, n'est jamais que l'assemblage d'un grand nombre d'événements élémentaires, alors les équations macroscopiques de la thermodynamique devraient pouvoir être déduites des équations microscopiques de la mécanique. Et s'il existe un lien très fort entre l'énoncé des lois microscopiques et ceux des lois macroscopiques, alors c'est une question de logique que les unes ne peuvent être réversibles si les autres ne le sont pas. Mais le lien entre les descriptions microscopiques et macroscopiques n'est pas aussi simple et direct que celui établi par une déduction (Chapitre 2). Ces descriptions se présentent généralement à des niveaux d'organisation différents et

d'après des théories différentes. Il peut en effet y avoir un décalage important qui peut prendre différentes formes : les concepts peuvent être hétérogènes en ce qu'ils n'ont pas la même signification, ni la même testabilité, ni les mêmes vertus cognitives. Il s'agit de raisons fréquemment invoquées afin d'affirmer l'échec d'une réduction épistémologique ou un cas d'émergence (Chapitre 2). D'ailleurs, plusieurs disciplines sont considérées comme autonomes, comme la thermodynamique et la biologie. Et tel que discuté, ce n'est pas parce qu'une loi est IRT que nécessairement l'observation des phénomènes qu'elle décrit en rendra compte de ce caractère d'IRT. Pour reprendre le même exemple, le comportement d'un corps en mouvement comme une boule de billard peut être décrit par une loi IRT comme celles de la mécanique, mais la description peut être celle, par exemple, d'un mouvement allant de gauche à droite, montrant ainsi une asymétrie dans les faits.

La mécanique statistique propose une solution au problème de l'asymétrie thermodynamique. Il a été vu (Chapitre 3) que l'entropie statistique pouvait se définir par le logarithme du volume en phase associé au macro-état du système et qu'elle pouvait être interprétée comme une mesure du nombre de micro-états compatibles avec ce macro-état et aussi, donc, comme la probabilité du macro-état. Or, comme la portion de l'espace des phases occupée par un état d'équilibre est beaucoup plus grande que toute autre portion de cet espace, il est en général soutenu, comme l'a fait Boltzmann, qu'il est très probable qu'un système dans un état quelconque tende vers l'état d'équilibre. Cette stratégie explicative soulève plusieurs questions, mais elle semble à tout le moins plausible (plus de détails au Chapitre 5). Un des problèmes d'importance auxquels elle fait face concerne l'objection à l'effet que, compte tenu des lois de la mécanique classique, si un système est dans un état de non-équilibre au temps t_2 , alors il *très probable* que

- (1) l'état de non-équilibre au temps t_2 évoluera vers un état d'équilibre au temps t_3 ,
- (2) l'état de non-équilibre au temps t_2 a évolué d'un état d'équilibre au temps t_1 ,

où $t_1 < t_2 < t_3$. Malgré tout, des phénomènes tels que (2) ne sont jamais observés. Si l'objectif initial est d'expliquer les phénomènes observés, cela semble plutôt raté. Le problème n'est pas tant que cette approche décrive des phénomènes qui ne se produisent pas (2), mais plutôt qu'elle fasse des prédictions qui soient falsifiées. Une théorie en contradiction avec les faits, comme en (2) et nonobstant les difficultés de vérifier cette contradiction dans le cas d'un énoncé probabiliste, est en effet problématique. Autrement dit, il faut trouver un moyen de conserver (1) et de rejeter (2).

Plusieurs solutions ont été proposées pour faire face à ce problème (voir Price 1996 et Callender 2011a pour une vue d'ensemble). La plus notoire est sans doute l'*hypothèse du passé* (« *Past-Hypothesis* »). Cette solution a recours aux *conditions initiales* où dans un passé très distant du présent, comme au début de l'univers, l'entropie de l'univers entier était très peu élevée et cet état de fait pourrait s'expliquer par

la théorie du Big Bang qui affirme qu'au début de l'univers, son volume était très petit et donc aurait atteint l'équilibre rapidement. Compte tenu du fait que *tous* les systèmes physiques proviennent de ces singulières conditions initiales, ils auraient tous à un moment donné eu une entropie plus faible qu'à un moment ultérieur. Cette solution vise donc à fournir une explication à l'asymétrie thermodynamique (Chapitre 5 ; Gal-Or 1972 ; Sklar 1993 ; Albert 2000 ; Callender 2011a).

En revanche, l'approche visant à faire de l'entropie un *explanans* des propriétés temporelles, soit l'interprétation ou théorie entropique de la directionnalité du temps, fait face à plusieurs problèmes. D'abord, si un système n'a pas d'entropie bien définie – car après tout ni la thermodynamique ni la mécanique statistique n'ont une prétention à s'appliquer à tous les systèmes possibles de l'univers – alors on peut se demander si la direction du temps est aussi indéterminée pour ce système, ou si le temps ne s'est pas arrêté, ou inversé, etc. Ensuite, cette tentative de réduction des propriétés temporelles aux propriétés entropiques ne peut être valide que si ces dernières sont jugées comme étant plus primitives. Mais si un phénomène est observé où il y a diminution spontanée d'entropie, une diminution tandis que le système est isolé de son environnement, alors il est fort à parier que l'on soit tenté de falsifier le principe entropique que de dire que le temps s'est inversé. Enfin, cette interprétation ne mène à aucune conséquence observable, car si, effectivement, il y avait diminution spontanée d'entropie, alors on pourrait affirmer que le temps s'est inversé, mais pas que l'interprétation est fausse. De sorte que cette théorie serait à l'abri de potentielles falsifications. De plus, et tel que discuté, l'inférence à l'effet que le temps, ou plutôt son cours, puisse être caractérisé par un processus particulier, donc localisé, ne semble pas justifiée. Car une diminution spontanée d'entropie, dans un endroit particulier de l'espace, n'aurait aucune conséquence directe sur le reste de l'univers, donc l'inversion du cours du temps, dans ce cas précis, n'aurait qu'une incidence locale.

Mais il est parfois soutenu, plus modestement, que l'asymétrie thermodynamique permettrait d'expliquer d'autres asymétries temporelles. Par exemple, elle expliquerait l'asymétrie de la mémoire (Reichenbach 1956), comme quoi on se souvient du passé et non du futur, ou encore celle de la causalité (Eckhardt 2006), comme quoi la cause précède toujours l'effet. Le principal problème avec ces approches est qu'elles tentent d'expliquer un phénomène jugé universel, comme la mémoire ou la causalité, à tout le moins ce qui est universel serait l'asymétrie dans les phénomènes instanciant ces concepts respectifs, alors que le concept d'entropie peut difficilement prétendre à une telle universalité, ne serait-ce qu'en raison du fait que la théorie qui le définit n'est pas universelle, et s'il n'est défini que pour les systèmes à l'équilibre, comme ce semble être le cas. À tout le moins sa généralisation est loin de faire l'objet de consensus.

Ce qui motive ces tentatives semble être une interprétation plutôt libérale du principe entropique, comme quoi il y aurait une tendance de passer d'un état peu probable à un état plus probable. Mais ce n'est certainement pas leur faire justice que de les amalgamer de la sorte. Pour y parvenir, il faudrait faire appel des outils de la mécanique statistique, lesquels seront présentés au prochain chapitre.

4.2 Thermodynamique

La thermodynamique est une discipline robuste et puissante en ce sens où elle est a été corroborée un nombre incalculable de fois et explique une très grande variété de phénomènes. Il n'empêche que plusieurs tentatives de réduction, d'explication et de réaxiomatisation (justification) ont été proposées, et que plusieurs études ont été consacrées à une relecture de ses fondements. La raison principale en est sans doute qu'il s'agit, tel que mentionné, d'une science dite « phénoménologique », en ce qu'elle s'appuie essentiellement sur des observations et qu'elle ne propose pas de modèle de la structure des objets observables. C'est ce qui en fait à la fois la force et la faiblesse, car elle s'accorde jusqu'à un certain point avec le sens commun mais s'éloigne aussi du schème conceptuel souvent prisé consistant à exprimer mathématiquement les phénomènes par des trajectoires de particules idéalisées. Mais la thermodynamique est aussi « l'esclave maltraité de plusieurs maîtres », des ingénieurs aux physiciens théoriciens en passant par les biologistes et les météorologistes. Tous n'ont pas la même compréhension ni les mêmes intérêts et c'est pourquoi elle est souvent en proie à de nombreuses modifications, voire « mutilations ».

En première approche, la *thermodynamique classique* ou orthodoxe peut être définie comme une science empirique reposant sur des « axiomes d'une validité empiriquement établie » (Lindsay & Margenau 1936 : 212), ou comme une science visant à « caractériser les corps physiques macroscopiques en termes de quantités macroscopiques observables » (Yoshioka 2007 : 3). Plus près de son histoire que de sa méthodologie, elle peut être décrite comme l'étude des phénomènes impliqués dans la production de travail au moyen de la chaleur, ou encore, de manière plus abstraite, de « l'interaction des transformations d'énergie thermique et mécanique » (Uffink 2001 : 309). Mais on a aussi vu (Chapitre 3) que la description de ces phénomènes a mené à l'élaboration du concept d'entropie, sorte de contrepartie du concept d'énergie, permettant de rendre compte des pertes d'énergie mécanique malgré la conservation d'énergie. En ce sens, la thermodynamique se définit comme « la science de l'énergie et de l'entropie » (Van Wylen, Sonntag & Desrochers 1992 : 16), ou simplement comme « la science des flux et transformations d'énergie » (Haddad, Chellaboina & Nersesov 2005 : 1). On constate donc que la thermodynamique se définit généralement comme une discipline scientifique décrivant des phénomènes dits macroscopiques et observables impliquant principalement l'énergie sous forme de chaleur et travail. Sa force réside ainsi en grande partie, on le rappelle, dans une base empirique solide et dans ce concept considéré comme inébranlable qui est celui de l'énergie.

La thermodynamique classique ne s'attarde donc pas à décrire la structure des systèmes et ne présente aucune hypothèse quant aux entités microscopiques pouvant constituer ces systèmes, qu'elle

traite pour l'essentiel comme des « boîtes noires ». Son principe empirique sans doute le plus fondamental est celui à l'effet qu'un système macroscopique, une fois laissé à lui-même, c'est-à-dire une fois suffisamment isolé de son environnement, atteint finalement un état d'équilibre et s'y maintient de telle sorte qu'aucun changement observable ne se produit. Mais cette théorie offre très peu d'outils pour décrire les processus car rien ne permet de décrire un état de non-équilibre. Seuls les processus commençant et se terminant dans un état d'équilibre, ou ceux se produisant très lentement, de sorte que chaque état de la séquence est à l'équilibre, sont expliqués par la théorie.

Plusieurs commentaires sont de règle suite à cette description. D'abord, le fait que la théorie s'abstienne de décrire la structure des systèmes lui permet d'avoir une généralité d'application qui a été mise à l'épreuve avec un certain succès dans des cas comme les gaz quantiques et les trous noirs. Mais comme les états de non-équilibre de même que les processus irréversibles constituent davantage la norme que l'exception, la prétention de la thermodynamique à expliquer divers phénomènes s'en trouve diminuée. Ainsi, le vocable « thermodynamique » semble inapproprié si aucun mécanisme n'est fourni sur le passage d'un état de non-équilibre vers un état d'équilibre et s'il n'y a pas vraiment d'équations différentielles par rapport au temps. En ce cas, il serait plus juste de parler de thermostatique. Cela ne signifie toutefois pas que la théorie n'a rien à voir avec le changement. Ensuite, comme les états d'équilibre sont par définition des états qui demeurent inchangés s'ils ne sont pas perturbés, tout processus doit donc être déclenché par une intervention extérieure, c'est-à-dire provenant de l'environnement du système à l'équilibre. En effet, que le système atteigne toujours l'équilibre après un certain temps implique qu'il peut ne pas être à l'équilibre et cet état de non-équilibre être reproduit même s'il n'est pas vraiment caractérisé par la théorie.

Des caractérisations différentes de la thermodynamique sont bien sûr disponibles. Elles s'appuient pour la plupart sur des développements récents, ou à tout le moins plus récents que ceux ayant formé l'approche orthodoxe de Clausius, Kelvin et Planck. La thermodynamique classique a par exemple influencé la chimie en s'intéressant au bilan énergétique des réactions chimiques, dans ce qui est convenu la *thermodynamique chimique*. C'est pourquoi on peut la caractériser d'après les changements d'énergie et de composition chimique (Thess 2011). Mais compte tenu de l'avènement de l'atomisme et de la mécanique statistique, qui propose en somme sa réduction à la mécanique, il n'est pas étonnant que la thermodynamique soit parfois définie comme une discipline décrivant des systèmes approximant le comportement stable de grands ensembles de particules. Par exemple, elle peut être décrite comme une discipline décrivant les propriétés statistiques moyennes des systèmes macroscopiques (Salem & Fröhlich 2006 : 1045), voire même comme la théorie qui « émerge » de la mécanique statistique à la limite des grands systèmes (Sethna 2011 : 114). Cependant, il n'est pas si évident que la thermodynamique puisse se réduire, dans un sens philosophiquement satisfaisant, à la mécanique, fut-

elle accompagnée du calcul des probabilités (une problématique largement discutée au Chapitre 5#). Ainsi, la question de savoir s'il est opportun de définir les concepts de la thermodynamique en fonction d'une théorie microscopique particulière demeure ouverte¹⁵.

Évidemment, plusieurs ouvrages ont été consacrés à la thermodynamique et à ses fondements. L'approche adoptée ici pour présenter l'entropie thermodynamique est dans un premier temps l'approche orthodoxe, celle de Clausius, Kelvin et Planck. D'autres approches ont bien sûr été développées. D'abord, celle de Gibbs (1902), qui ne s'attarde pas à trouver une correspondance empirique pour les variables d'états. Il y a aussi celle de la *thermodynamique axiomatique*, qui se veut bien souvent iconoclaste, en ce qu'elle propose de nouvelles bases, de nouveaux axiomes, souvent éloignés des généralisations empiriques initiales (par ex. Lieb & Yngvason 1999 et 2003). Mais il a été vu au chapitre 3# que le besoin de généraliser la thermodynamique *qua* thermostatique aux phénomènes irréversibles et aux états de non-équilibre s'est vite manifesté. La *thermodynamique des processus irréversibles* (par ex. Prigogine 1968 ; Yourgrau, Merwe & Raw 1966), qui constitue en quelque sorte l'orthodoxie de la thermodynamique du non-équilibre, tente ainsi d'étendre la portée de la thermodynamique aux systèmes dans des états de non-équilibre où, typiquement, les quantités thermodynamiques sont représentées par des champs variables continus à l'instar de ce que l'on retrouve en mécanique des milieux continus. L'approche développée par la *thermodynamique rationnelle* n'est guère différente. Leur différence réside principalement dans leur prétention en ce que cette dernière n'a jamais été limitée à une théorie linéaire¹⁶. Enfin, la *thermodynamique prolongée* aurait la prétention pour sa part de répondre aux objections adressées à celles-ci. Malgré ces développements, ces théories ne modifient pas radicalement le concept d'entropie, bien qu'elles l'utilisent différemment.

Enfin, notons aussi la *thermodynamique statistique*, laquelle rejette une hypothèse fondamentale de l'approche orthodoxe, à savoir qu'un état d'équilibre ne serait pas un état dont les quantités physiques sont constantes, mais serait plutôt soumises à des fluctuations (voir Le Bellac, Mortessagne & Batrouni 2004). Puis, il y a la *thermodynamique relativiste* qui applique les conclusions de la théorie de la relativité à la thermodynamique, comme celle voulant que les corps chaud soient plus lourds que les corps froids

¹⁵ Salem & Fröhlich (2007 : 1046) soutiennent que la thermodynamique est une théorie « indépendante » (« *self-contained* ») mais affirment du même souffle qu'il est intéressant de dériver ses lois fondamentales à partir d'une description *ab initio* microscopique, comme la mécanique quantique statistique. S'il est possible et intéressant de procéder à une telle dérivation, alors en quoi la thermodynamique est vraiment « indépendante » ?

¹⁶ La « tragico-comédie » de la thermodynamique, pour reprendre l'expression de Truesdell (1980), se poursuit encore aujourd'hui. En effet, Müller (1972), un ancien étudiant de Truesdell, a soulevé une objection à l'effet qu'un principe fondamental de la thermodynamique rationnelle (« principe d'invariance référentiel matériel ») développée par ce dernier, ne tenait pas. Truesdell (1976) a répondu à cette objection, mais Müller (2007) affirme qu'il l'a plutôt « ignorée » (éludée). Alors que Müller (2007) signale le côté « humoristique » de l'article de Rivlin (1970), les éditeurs de la revue fondée par Truesdell lui-même y voient plutôt un article « malheureux » et « une attaque à peine voilée » contre ce dernier (Ball & James 2002).

car leurs atomes ont une plus grande vitesse donc une grande masse, et celle à l'effet qu'aucune influence ne peut se propager plus vite que la lumière. Elle généralise ainsi la thermodynamique aux conditions extrêmes, comme des températures extraordinairement élevées, bien que la température au centre du soleil ne soit *pas* suffisante pour qu'il vaille la peine d'appliquer les corrections relativistes (voir Tolman 1934 ; Hayward 1999 ; Müller 2007).

Toutes ces disciplines peuvent être regroupées dans ce que l'on peut appeler la *thermodynamique générique*. Mais leur tronc commun est difficilement identifiable. On peut néanmoins dire qu'il s'agit d'une *théorie de champ* (*field theory*), dont l'objectif est la détermination de différents champs (densité de masse, vitesse d'écoulement et température) et qui s'appuie sur le premier et le second principes de la thermodynamique. En introduisant la notion de champ, postérieure à ces deux principes, elle dresse donc des ponts entre le passé et le présent de la discipline.

Dans cette section, je souhaite d'abord présenter les principaux paramètres (états et processus) et principes de la théorie thermodynamique. Je vais ensuite discuter de la signification de son second principe tel qu'il origine des travaux de Carnot, Kelvin et Clausius. La fonction d'état introduisant le concept d'entropie sera ensuite présentée et discutée, de même que le principe entropique qui en découle. Je vais aussi présenter, brièvement, des tentatives d'extension de ce principe aux phénomènes de non-équilibre.

4.2.1 États et processus en thermodynamique

Il a été vu que le nom « thermodynamique » peut sembler inapproprié puisque la thermodynamique classique n'a à peu près rien de « dynamique » en ce qu'elle se concentre sur les états d'équilibre, et qu'en ce sens, il serait plus juste de parler de « thermostatique ». En effet, bien qu'elle fasse largement référence à des « processus », aucune description de séquence d'états dans le temps n'est fournie. Plutôt, les processus sont introduits afin de déterminer les états finaux possibles à partir de certains états initiaux. Mais ce qui se passe entre ces états n'est à peu près pas explicité.

Un *état thermodynamique* est un état caractérisé par des variables *macroscopiques* et une situation d'*équilibre*. Ces variables peuvent être *intensives*, c'est-à-dire indépendantes de la taille ou de la masse du système, ou *extensives*, c'est-à-dire qu'elles dépendent de la taille du système. Un état est à l'équilibre si ses propriétés sont suffisamment stables dans le temps, où chaque variable intensive est définie et a la même valeur en tout point du système, de sorte que l'état soit indépendant du temps et donc de l'histoire du système. Et un état est dit macroscopique si ses propriétés peuvent être mesurées et décrites par un petit nombre de quantités macroscopiques, comme la pression P , la température T et le volume

V , sans égard aux fluctuations possibles¹⁷. Une chaudière d'une locomotive en train d'exploser n'est *pas* un exemple de système dans un état thermodynamique. La caractérisation des états thermodynamiques par rapport à l'équilibre et à la macroscopicité constitue le principal critère distinguant les variables thermodynamiques de celles qui ne le sont pas (Landsberg 1956 : 365). En effet, la thermodynamique considère la spécification de l'état dynamique de l'ensemble des constituants microscopiques comme « superflue » (Fermi 1936 : 1). En conséquence, un état thermodynamique est un état caractérisé par un petit nombre de variables dont les valeurs sont constantes et uniformes, de sorte que leurs fluctuations ainsi que leurs variations spatiales sont négligées¹⁸.

Les quantités macroscopiques caractérisant un système thermodynamique constitué d'une substance quelconque, comme l'eau, dont la masse est fixe ne sont pas des variables indépendantes et sont liées par une relation de la forme $f(P, T, V) = 0$, appelée *fonction d'état* ; cette fonction est parfois exprimée aussi en fonction du nombre de moles d'une substance ou de son potentiel chimique. La forme de cette relation dépend des propriétés particulières de la substance. Chacune des trois variables d'une relation de ce type peut être exprimée par une fonction des deux autres en résolvant l'équation précédente par rapport à une variable donnée. En conséquence l'état du système est complètement déterminé par au moins deux de ces trois variables P , T et V , à condition que le système soit à l'équilibre. Par ailleurs, un état est *stationnaire* si les paramètres qui le décrivent – comme la température, la pression ou l'entropie – ne dépendent pas du temps, en ce sens où les flux de différentes quantités physiques se compensent sans pour autant être nuls, ce qui le distingue donc d'un état d'équilibre.

Il est utile et même nécessaire de caractériser les systèmes thermodynamiques d'après les échanges d'énergie (chaleur et travail) et de matière à travers leurs frontières : (i) les systèmes *isolés* n'échangent ni énergie ni matière ; (ii) les systèmes *fermés* échangent de l'énergie mais sans matière ; on peut aussi distinguer les *systèmes fermés adiabatiques* qui échangent de l'énergie sous forme de travail mais pas sous forme de chaleur ; (iii) les systèmes *ouverts* échangent matière et énergie avec l'extérieur.

Les *processus thermodynamiques* se présentent principalement selon quatre types : (i) les processus *adiabatiques* où il n'y a pas d'échange d'énergie sous forme de chaleur entre le système et son environnement (ces processus concernent donc les systèmes isolés et les systèmes fermés adiabatiques) ;

¹⁷ Supposant qu'un système thermodynamique soit composé d'un grand nombre N de molécules, on peut facilement montrer que ses fluctuations sont proportionnelles à $1/\sqrt{N}$ et donc elles tendent vers 0 lorsque $N \rightarrow \infty$. Truesdell (1980 : 82) soutient que la possibilité d'identifier un « état » par un nombre fini de variables est une « illusion ». Sans doute a-t-il en tête que chaque corps macroscopique possède une infinité de degrés de liberté si l'on tient compte, par ex., de la position spatiale de chacune des molécules. Mais il faut savoir écarter ce qui n'est pas pertinent dans ce contexte, écarter le « bruit explicatif ».

¹⁸ Truesdell (1980), avec sa verve proverbiale, affirme que le terme « état thermodynamique » est ce « *king cobra in the pit of thermodynamic vipers* », et s'abstient ainsi d'utiliser ce terme qui ne représenterait aucun « concept rationnel ».

(ii) les processus *quasi-statiques* qui sont constitués d'une séquence lente d'états d'équilibre ; (iii) les processus *réversibles* où il est possible de renverser la séquence des états de ce processus, et un processus réversible est quasi-statique, et vice-versa) ; (iv) les processus *irréversibles* où il n'est *pas* possible de renverser la séquence des états de ce processus, et un processus irréversible n'est *pas* quasi-statique, donc il n'est *pas* à l'équilibre, et vice-versa.

4.2.2 Principes fondamentaux

La thermodynamique moderne est conçue autour de quatre principes fondamentaux¹⁹. Leur portée est impressionnante en ce qu'ils rendent compte d'une grande diversité de phénomènes²⁰. Le principe zéro, qui a atteint ce statut après la formulation et la cristallisation des deux principes suivants, introduit le concept de température, le premier principe celui d'énergie et le second principe celui d'entropie. Le troisième principe opère un chevauchement des descriptions proprement thermodynamiques, empiriques, phénoménologiques et macroscopiques avec les descriptions quantiques et microscopiques et détermine une valeur absolue à l'entropie. Du point de vue de la thermodynamique classique, ces principes sont des hypothèses fondamentales indépendantes dont la justification repose essentiellement sur leur adéquation empirique mais aussi théorique, puisqu'ils s'accordent autant avec l'expérience, dont ils se veulent une sorte de résumé, qu'avec d'autres principes scientifiques établis. Mais ils peuvent aussi recevoir, dans un autre contexte, une justification à partir d'hypothèses de nature statistique et/ou quantique sur la constitution de la matière formant les systèmes auxquels réfèrent les justifications empiriques. (Pour des discussions plus détaillées, voir Fermi 1936 ; Bridgman 1941 ; Guggenheim 1949 ; Desrochers, Sonntag & van Wylen 1992 ; Dugdale 1996 ; Salem & Fröhlich 2007 ; Çengel & Boles 2006 ; Atkins 2007 ; Kondepudi 2008.)

1. Le *principe zéro de la thermodynamique* exprime la transitivité de l'équilibre et s'énonce comme suit : si deux systèmes sont à l'équilibre thermique avec un tiers système, alors ils sont tous à l'équilibre thermique. Ce principe semble *prima facies* trivial, mais il ne peut être déduit d'autres principes et précède les premier et second principes de la thermodynamique²¹. Il suppose d'une part l'existence de

¹⁹ J'ai préféré le terme « principe » à celui de « loi » en raison des multiples formulations qu'ils recevaient et qui ne sont pas toujours mathématisées. De plus, dans le cadre d'une entreprise critique, le mot « principe » se présente, à mon sens, avec moins de connotations épistémologiques que celui de « loi ». On trouve les deux termes dans la littérature, quoique le terme « law » (« *Second Law of Thermodynamics* ») est définitivement privilégié en anglais.

²⁰ Le titre de Atkins (2007), *Four laws that drive the universe*, est à cet égard probant.

²¹ Il est toutefois possible, en mécanique statistique, d'offrir une dérivation du principe zéro de la thermodynamique en définissant la température à partir de l'entropie, vue comme une quantité non-décroissante (voir Penrose 1979 : 156).

systèmes thermodynamiques approchant un état d'équilibre lorsqu'ils sont isolés, qu'on appelle *réservoirs thermiques*, et d'autre part l'existence de cloisons *diathermiques* permettant un transfert de chaleur et des changements éventuels à l'intérieur de ces cloisons. Toutefois, la chaleur n'a pas besoin pour l'instant d'être définie : si aucun changement ne survient alors que deux systèmes sont mis en contact l'un avec l'autre, alors ces deux systèmes sont considérés à l'équilibre thermodynamique. À l'instar de la pression qui est l'indicateur universel de l'équilibre mécanique, la *température* est l'indicateur universel de l'équilibre thermique. C'est-à-dire, si deux systèmes sont à l'équilibre thermique, alors ils ont une propriété en commun, celle d'avoir la même température. Ce principe exprime donc la transitivité des réservoirs thermiques et circonscrit l'existence des thermomètres permettant de mesurer la température, par un changement observable, comme l'expansion d'un volume de mercure par exemple.

2. Le *premier principe de la thermodynamique* introduit le concept d'énergie. Ce concept fondamental est défini à partir du principe de conservation de l'énergie qui stipule que l'énergie de tout système isolé de son environnement demeure constante. Lorsqu'un système interagit avec son environnement et passe d'un état A à un état B , son énergie (U) peut changer d'une certaine quantité $dU = U_B - U_A$. Il y a bien évidemment plusieurs façons d'intervenir sur un système, comme il y a plusieurs façons de faire passer un système d'un état A à un état B . Les deux types généraux d'intervention permettant de modifier un système et de changer son énergie sont la chaleur (Q) et le travail (W). La chaleur peut se définir comme la forme d'énergie qui est transférée entre deux systèmes (ou un système et son environnement) en présence d'une différence de température, tandis que le travail peut se définir comme la forme d'énergie qui est transférée par l'action d'une force sur une distance. Si ce n'est pas de la chaleur, alors il s'agit du travail, et vice-versa.

Le premier principe ne fait que définir la variation de l'énergie en équivalence avec des quantités de chaleur et de travail²². Il s'exprime comme suit :

$$dU = \delta Q + \delta W.$$

Ainsi, selon ce principe, lors d'un processus cyclique où le système revient à son état initial et donc à un même niveau d'énergie, la variation d'énergie est nulle ($dU = 0$) et par conséquent la quantité de travail

²² En pratique, la distinction entre la chaleur et le travail semble évidente ; on ne saurait confondre le déplacement d'un poids avec l'élévation de température. En théorie, par contre, une distinction claire et rigoureuse est moins simple qu'il n'y paraît. Par exemple, Gould & Tobochnik (2010 : 18) affirment que la chaleur consiste en un transfert d'énergie de manière « incohérente » tandis que qu'un travail est un transfert d'énergie de manière « cohérente ». En revanche, Penrose (1970 : 158) soutient que lors d'un transfert d'énergie sous forme de travail peu de « degrés de liberté observables » sont impliqués tandis que lors d'un transfert d'énergie sous forme de chaleur plusieurs « degrés de liberté moléculaires inobservables » sont impliqués, de sorte que la distinction entre ces deux concepts reposerait sur la possibilité de mesurer expérimentalement certains paramètres affectant l'énergie globale (hamiltonien) du système. (Voir Chapitre 5.)

fournie par le système est égale à la chaleur fournie au système ($-\delta W = \delta Q$). Puisque la quantité de travail ou de chaleur extraite d'un système dépend du processus utilisé, ni Q ni W ne peuvent caractériser entièrement un changement d'état et ne peuvent donc pas être qualifiés de variables d'état (et ne peuvent donc pas être représentés par une différentielle exacte, d'où le terme δ). Il est alors plus juste de dire que la chaleur et le travail ne sont pas à proprement parler des formes d'énergie, mais plutôt des façons ou des méthodes particulières d'échanger ou de transférer de l'énergie. Autrement dit, chaleur et travail sont de l'énergie en transition, deux types d'interactions énergétiques entre un système et son environnement, qui exigent, pour une description complète, que soient spécifiées leur magnitude et leur direction (on parle alors de « quantités directionnelles »).

3. Brièvement, le *second principe de la thermodynamique* suppose aussi la définition et donc la distinction des concepts de chaleur et de travail. Il stipule en substance qu'une quantité de chaleur donnée ne peut être entièrement convertie en travail, bien que l'inverse soit possible. Il peut aussi faire être formulé en fonction du concept d'entropie, ce qu'on appelle le principe entropique. (Voir la prochaine section.)

4. Le *troisième principe de la thermodynamique* stipule que l'entropie tend vers zéro lorsque la température tend vers le zéro absolu. À l'instar du premier et du second principe s'appuyant sur des constatations empiriques à l'effet que certains processus étaient impossibles à obtenir, la version phénoménologique du troisième principe stipule qu'il n'y a aucune séquence finie de processus cycliques permettant de refroidir un corps au zéro absolu. Sa version microscopique stipule que l'entropie d'une structure « parfaitement cristalline » est nulle au zéro absolu²³. Mais ce principe n'introduit pas de nouvelle fonction et en ce sens il est différent des trois autres principes de la thermodynamique.

5. Lorsqu'un système isolé est dans un état d'équilibre, en raison des trois premiers principes, on peut remplacer δQ par TdS et δW par PdV , de sorte que

$$dU = TdS - PdV, \text{ ou } TdS = dU + PdV$$

(Bejan 2006 ; Çengel & Boles 2006). Il s'agit de l'*équation fondamentale de la thermodynamique* dont les conséquences par rapport à des applications à une grande variété de systèmes et de situations constituent le « point focal » de la thermodynamique (Dugdale 1996 ; Müller 2003)²⁴.

²³ Les considérations microscopiques impliquent des subtilités qui seront discutées au prochain chapitre. Par ailleurs, comme l'entropie est une fonction d'état elle dépend bien sûr des caractéristiques de l'état du système, comme la température et la pression ; mais au zéro absolu, l'entropie est indépendante de la pression.

²⁴ Pour Müller (2003 : 21), son importance ne peut être surestimée et elle vaudrait littéralement des milliards à l'industrie chimique.

4.2.3 Le second principe de la thermodynamique

Le premier principe de la thermodynamique exprime l'impossibilité de créer ou d'annihiler de l'énergie. Cela équivaut à l'impossibilité du mouvement perpétuel de premier type, qui est défini comme un mouvement obtenu sans dépense d'énergie. Autrement dit, il y a une égalité quantitative entre les différentes formes d'énergie, entre la chaleur et le travail. Mais le premier principe ne précise pas sa portée car il ne place aucune contrainte sur les transformations d'énergie d'une forme à l'autre. Il est en effet possible de convertir une quantité arbitraire de travail en chaleur, par le truchement de la friction par exemple. Il y a cependant des limites, pour un système donné, à la conversion ou transformation de la chaleur en travail. En d'autres mots, l'égalité quantitative entre le travail et la chaleur tient dans un sens mais pas dans l'autre. Si ce n'était pas le cas, il serait possible de construire une machine qui ne ferait que transformer la chaleur des objets ambiants en travail. Et puisque la chaleur est pour ainsi dire illimitée, cela reviendrait à construire une machine à mouvement perpétuel de second type, soit une conversion complète de la chaleur en travail. Or le second principe élimine cette possibilité.

Ainsi, comme le souligne Planck (1945 : 79), en se limitant à cette égalité quantitative, le premier principe est insuffisant à une caractérisation univoque des processus naturels. En clair, cela signifie que le premier principe seul ne permet pas de faire des prédictions valides, des descriptions adéquates ou des choix judicieux (pour modifier notre environnement à notre avantage) à propos d'une grande variété de phénomènes naturels. En voici quelques exemples. D'abord, si l'énergie se conserve, alors pourquoi parle-t-on de *consommation* d'énergie et d'*efficacité énergétique* ? Si elle se conserve, elle ne peut pas être consommée, puisqu'elle ne se perd pas, et donc l'efficacité devrait être toujours optimale et égale à l'unité, et donc il n'y aurait pas de processus plus efficace qu'un autre. Aussi, pourquoi l'énergie a-t-elle un *coût* si elle se conserve ? L'énergie que je mets dans ma voiture sous forme de carburant ne peut pas se perdre, alors pourquoi ne puis-je pas revendre la chaleur de mon moteur ? Et pourquoi deux systèmes ayant la même énergie ne semblent-ils pas présenter la même « utilité » ou la même « exploitabilité » ? Par exemple, une batterie et une chaudière d'eau chaude peuvent posséder chacune la même quantité d'énergie, mais je sais, du moins intuitivement, qu'avec l'une je puis faire démarrer une voiture ou faire fonctionner mon ordinateur, tandis qu'avec la seconde je suis tout juste bon à faire la vaisselle. Le second principe aurait la prétention d'unifier les réponses à ces questions au sein d'une même stratégie explicative.

La force ou la robustesse du second principe repose sur le « fait physique » de l'irréversibilité. Mais cette affirmation ne va pas de soi. Bien que ce concept puisse être compris intuitivement dans certaines situations simples, il a été vu précédemment qu'il pouvait recevoir plus d'une interprétation et que son

application n'était pas aussi évidente qu'on voudrait le croire. Si l'on en croit certains auteurs (par ex. Uffink 2001), son articulation au sein du second principe ne serait pas assurée et c'est l'une des raisons de la variété de formulations de ce principe. Une autre raison se trouve sans doute dans la tension que présente l'irréversibilité que se propose d'encadrer le second principe face à la réversibilité temporelle des équations de la plupart des théories scientifiques, au premier chef la mécanique newtonienne. Les deux formulations les plus célèbres de ce principe sont celles de Clausius²⁵ et Kelvin²⁶ (Chapitre 3). Mais le théorème de Carnot²⁷ mérite qu'on s'y attarde car il a su inspirer ces deux auteurs et ouvrir les débats sur la notion d'irréversibilité (*ibid.*). L'examen le plus exhaustif, et peut-être aussi le meilleur, du second principe de la thermodynamique et du principe entropique, est sans doute celui de Uffink (2001). C'est pourquoi la discussion qui suit s'y référera afin de répondre à certains de ses arguments. Il est important de noter, encore une fois, que le second principe de la thermodynamique, qui fait appel essentiellement aux notions de chaleur et de travail, sera distingué du principe entropique (discuté plus loin), qui implique bien sûr le concept d'entropie, car les problèmes épistémologiques que ces deux principes rencontrent ne sont pas nécessairement coextensifs.

L'objectif ici n'est pas tant de rendre justice aux acteurs historiques (voir Chapitre 3) mais plutôt d'établir de manière claire la signification du second principe de la thermodynamique. Cela peut et doit se faire à la lumière de ce qui a déjà été fait dans cette voie. Car on ne saurait remettre en cause l'idée qu'il y ait quelque chose de fondamental dans ce principe – comme l'impossibilité du mouvement perpétuel – et il est plutôt question de savoir de quoi il s'agit. Il est maintenant commun d'entendre des commentaires sur le ton de la plainte à l'effet que tout le monde sait bien ce que le second principe signifie, mais que personne ne s'accorde vraiment sur sa signification. Je propose ainsi une « reconstruction » s'appuyant sur les développements fondamentaux ayant mené aux premières formulations du second principe en limitant autant que faire se peut les hypothèses, que ce soit sur la nature de la chaleur, les processus quasi-réversibles ou les réservoirs infinis de chaleur. Cette reconstruction tentera de dégager l'« essence » de ce principe, c'est-à-dire le point commun indispensable aux différentes formulations et ce, afin de les mettre en lumière et ainsi offrir un plan d'analyse au concept d'entropie. Car s'il est l'objet de critiques malgré qu'il soit fondamental, bien souvent c'est en raison de desiderata épistémologiques très généraux censés encadrer l'élaboration des théories scientifiques.

²⁵ La chaleur ne peut passer d'elle-même d'un corps froid à un corps plus chaud.

²⁶ Si une machine est telle que, lorsque son fonctionnement est renversé, ses éléments physiques et mécaniques sont tous renversés, alors elle produit autant de travail qu'il est possible d'en produire à partir d'une quantité de chaleur donnée.

²⁷ Les machines réversibles présentent un rendement maximal : pour une quantité de chaleur (« calorique ») donnée, le travail (« puissance motrice ») est optimal.

4.2.3.1 RECONSTRUCTION

La première étape de cette reconstruction remonte aux travaux de Fourier et de la loi qui porte son nom, laquelle stipule que le flux de chaleur entre deux corps est proportionnel à la différence de température entre ces corps et qu'il se déplace dans la direction de la température décroissante. Autrement dit, une différence de température est nécessaire à un transfert de chaleur, lequel se fait du chaud vers le froid²⁸. La seconde étape concerne l'interconvertibilité entre chaleur et travail, qui est un cas particulier du principe de conservation de l'énergie et une version simplifiée du premier principe de la thermodynamique. En effet, cette interconvertibilité découle logiquement du principe plus fort qu'est l'équivalence entre chaleur et travail, laquelle a été pressentie par Carnot et prouvée expérimentalement par Joule. En termes simples, elle stipule que la chaleur peut être transformée en travail et vice-versa. Mais les travaux de Carnot rendent clair que la conversion de la chaleur en travail ne peut être complète et qu'elle est proportionnelle à la différence de température, ce qui constitue la troisième et dernière étape. C'est-à-dire que le rendement d'un processus, qui se définit par le ratio du travail obtenu divisé par la chaleur fournie, est strictement inférieur à l'unité. Cela signifie que la conversion complète d'une quantité de travail en chaleur est possible, tandis que l'opération inverse ne l'est pas. En somme, il est nécessaire d'avoir une différence de température (ΔT) pour obtenir un flux de chaleur (Q) et donc un travail (W). En d'autres mots, sans différence de température, pas de flux de chaleur et pas de production de travail. De manière plus schématique, cette séquence peut s'exprimer comme suit :

SÉQUENCE THERMODYNAMIQUE (ST). $\Delta T \rightarrow Q \rightarrow W [= (1 - T_-/T_+)Q = Q_+ - Q_-]$.

La quantité de travail *maximale* (W) pouvant être obtenue est strictement inférieure à la quantité de chaleur transférée (Q), car elle dépend des températures d'opération (T_- et T_+) selon la formule du rendement de Carnot, dont la valeur est strictement inférieure à l'unité (rappel : $\eta = 1 - T_-/T_+ < 1$)²⁹. La ST a donc comme corollaire l'impossibilité d'un mouvement perpétuel de second type, soit un processus capable de convertir entièrement la chaleur en travail³⁰.

²⁸ On peut ignorer ici le transfert thermique sous forme radiative et le cas particulier (quantique) d'une température négative (voir la note un peu plus loin).

²⁹ Cet énoncé reprend différemment les slogans souvent rapportés aux étudiants de thermodynamique : le premier principe garantit l'équivalence entre le travail et la chaleur, de sorte qu'on ne puisse obtenir plus de travail d'un système que la quantité de chaleur fournie, donc « on ne peut pas gagner ! » (« *you cannot win!* ») ; tandis que le second principe garantit que le travail obtenu est inférieur à la quantité de chaleur fournie, donc « on ne peut même pas atteindre l'égalité ! » (« *you cannot break even!* »).

³⁰ Plusieurs exemples sont proposés afin d'illustrer cette impossibilité, où le caractère absurde qu'ils suscitent face à nos convictions est censé servir de preuve, à tout le moins temporairement. Si cette impossibilité était levée, alors il serait possible de faire accélérer un bateau en tirant la chaleur de l'océan ou de n'importe quel cours d'eau jusqu'au zéro absolu ; ou encore,

Cette séquence (logique au sens où elle présente une série de conditionnels) repose sur très peu d'hypothèses, car elle ne suppose rien de la constitution de la matière et ne s'appuie pas sur une théorie de la chaleur particulière. Elle n'est pas non plus liée à une caractérisation spécifique des états et des processus thermodynamiques, tel que discuté précédemment, ni, en particulier, aux processus cycliques. Elle se veut simple et générale : elle est donc compatible avec un certain nombre d'hypothèses auxiliaires qui ne sont pas explicitées, et elle ne devrait pas porter à polémique étant donné que ses hypothèses sont empiriquement bien établies, pour autant que les concepts de travail et de chaleur le soient. En effet, son affirmation principale concerne l'impossibilité d'un mouvement perpétuel de second type. Il semble que seule l'expérience permettrait d'écarter ce principe bien établi. Mais cela ne signifie pas que la ST a les moyens de ses ambitions, soit représenter l'essence du second principe, et c'est l'objectif de la discussion qui suit d'y répondre.

La ST peut être interprétée comme une « chaîne causale » dans le contexte d'un système thermodynamique où les concepts de température, de chaleur et de travail sont bien définis. Elle peut être interprétée comme une série de conditions nécessaires mais aussi suffisantes ou comme une série de manipulations permettant d'obtenir le résultat final selon le point de départ choisi, soit la différence de température ou l'investissement de travail. Cette interprétation est motivée par la portée générale recherchée visant à limiter encore une fois les hypothèses dont la validité pourrait être remise en cause. En décrivant une série comme étant « causale », dont les relations sont représentées par ' \rightarrow ', le mode de détermination, lequel peut bien sûr être décrit par une équation mathématique spécifiant l'évolution de la variable, demeure largement non spécifié. Car il n'est pas question d'harnacher la ST à une conception spécifique de la causalité. Ceci dit, la conception manipulationniste satisfait ce caractère général et intuitif de la production d'une séquence d'événements successifs (pour plus de détails, voir Pearl 2000 et Woodward 2003 ; aussi, Wallace 2014 propose de « réduire » la thermodynamique à une « théorie du contrôle »). En effet, une telle séquence ne s'appuie que sur des paramètres observables et peut être vérifiée expérimentalement, et la thermodynamique en fait état. En ce sens, la ST respecte certains desiderata empiristes « modestes » : elle peut être incomplète (en ce sens où un modèle de la constitution de la matière est absent), mais elle est robuste. Cette caractérisation est donc suffisante pour notre discussion.

Il importe d'abord de préciser encore une fois que cette séquence peut être inversée. C'est-à-dire qu'il est possible d'utiliser une quantité de travail pour obtenir une différence de température entre deux corps. Plus précisément, par le truchement d'un *autre* corps présentant des caractéristiques

le processus consistant à faire ralentir un corps par friction pourrait être inversé, et ce corps pourrait alors être accéléré en tirant la chaleur qu'il a cédée par friction.

particulières et d'une quantité de travail³¹, il est possible d'opérer un transfert de chaleur du corps froid vers cet autre corps plus froid, puis, suite à quelques opérations sur cet autre corps consommant la quantité de travail investie, d'opérer un transfert de chaleur de cet autre corps vers le corps chaud, alors moins chaud que cet autre corps. En définitive, donc, le transfert *direct* (contact de deux corps) de chaleur est *toujours* du chaud vers le froid. Ainsi, une différence de température produit un flux de chaleur entre deux corps, du chaud vers le froid, et ce sans intervention extérieure d'un autre système, tandis qu'un investissement de travail provenant d'un troisième système peut produire un flux de chaleur entre deux corps, mais cette fois du froid vers le chaud et de manière indirecte par le truchement d'un autre corps, créant ainsi une différence de température ou encore permettant de l'intensifier. Il y a donc une *asymétrie* dans le transfert de chaleur. De fait, alors qu'un transfert de chaleur du chaud vers le froid n'exige aucune dépense de travail, un transfert de chaleur du froid vers le chaud, au demeurant indirect puisqu'il requiert l'intervention d'un autre corps, demande une dépense de travail provenant de l'extérieur du système. Cette asymétrie ne peut être rompue car un transfert de chaleur inversé est impossible dans les mêmes conditions qu'un transfert de chaleur spontané, sans l'intervention d'un troisième corps.

Cette possibilité d'inverser le transfert de chaleur dans des conditions particulières est théorique, puisque la thermodynamique l'autorise, mais aussi pratique ou technique, car elle est effectivement réalisée avec les pompes à chaleur ou les réfrigérateurs, par exemple. Mais cette possibilité de réversibilité mérite examen.

4.2.3.2 DISCUSSION

D'abord, est-ce une réversibilité au sens d'IRT ? On a vu plus tôt qu'il s'agit d'une condition restrictive pour les processus mais qu'elle était souvent rencontrée dans les théories (et c'est pourquoi il semble bien qu'il faille trouver une explication à l'irréversibilité des phénomènes si ces théories sont justement censées les décrire). Il faut d'abord noter que les étapes de la ST ne peuvent être scindées de façon à démarrer la séquence à une étape arbitraire. En effet, si une différence de température donne lieu à un transfert de chaleur du corps chaud au corps froid, c'est-à-dire dans la direction de température décroissante, l'inverse est impossible, c'est-à-dire qu'il n'y a pas de transfert inverse de la chaleur du corps froid au corps chaud et ce, dans les mêmes circonstances. Autrement dit, si le système est défini uniquement par cet assemblage d'un corps chaud et d'un corps froid, alors le processus inverse est strictement impossible. Ce qui permet d'établir les processus possibles, en thermodynamique comme

³¹ Par exemple, du fréon (l'autre corps) et une quantité d'électricité (la quantité de travail) dans le cas d'un réfrigérateur électrique.

dans la plupart des théories scientifiques, exige la définition et l'identification des systèmes en question et de leur environnement. En mécanique statistique, cela revient à définir l'espace des phases, qui encadre les états possibles d'un système (voir le prochain chapitre). Car ce qui est possible pour un système isolé de son environnement ne correspond pas à ce qui est possible pour un système en interaction avec son environnement. Si le système considéré pour un certain processus n'est pas le même que celui du processus inverse, alors le critère de réversibilité au sens d'IRT ne peut être satisfait.

Uffink (2001) et Lervig (1972) concluent que la théorie de Carnot est non-IRT car, bien qu'elle interdise la production de travail sans consommation de ressources d'aucune sorte, ce qui reviendrait à produire un mouvement perpétuel, elle n'aurait en revanche aucun « scrupule » à propos de processus consommant une quantité de travail sans autre effet, ce qui définit un mouvement perpétuel de premier type inversé. Or, la ST est très semblable à la théorie de Carnot. Mais il est loin d'être évident que celle-ci puisse se résumer à la description qu'en font Uffink et Lervig. D'une part, la théorie de Carnot ne parle pas de processus consommant une quantité de travail sans autre effet. Or on ne saurait présumer ce qui n'est pas spécifié. Ce n'est pas parce qu'une théorie est muette par rapport à un certain phénomène qu'elle l'autorise, avec ou sans « scrupule ». De plus, la condition « sans autre effet » n'est pas claire puisqu'elle ne spécifie pas si un éventuel effet doit être situé dans le système ou dans son environnement. Mais dans les deux cas il y a bien un effet. Car, d'une part, un travail effectué sur un système produit un effet sur celui-ci, bien que l'état initial du système puisse être recouvré. Toutefois, et d'autre part, la production d'un travail a aussi un effet sur l'environnement puisqu'elle exige la consommation de ressources. Il n'y a en effet aucun problème, en principe, à ce qu'un mouvement soit soutenu indéfiniment si on lui fournit indéfiniment de l'énergie sous forme de travail.

Aussi, si le travail est obtenu d'une différence de température, comme c'est le cas chez Carnot et pour la ST, alors, dans les mêmes circonstances (à quelques ajustements près), la consommation de travail va créer une différence de température. De sorte que les deux théories seraient IRT en ce que la séquence des états peut être produite dans un sens comme dans l'autre pour un système donné. Mais il faut bien voir que la ST rejoint la théorie de Carnot dans le *cas limite* où la quantité de travail *maximale* est obtenue, autrement dit lorsque le rendement est maximal et décrit par la formule de Carnot. Une telle situation correspond chez Carnot à un processus réversible, tandis que le processus n'est pas spécifié dans la ST. Elles sont donc IRT seulement dans le cas limite d'un rendement maximal et non-IRT dans les autres cas.

Ensuite, est-ce que la théorie de Carnot et la ST sont irréversibles au sens de l'irrécupérabilité ? Dans une situation où le processus ne possède pas un rendement maximal, la quantité de travail produite à partir d'une quantité de chaleur est bien sûr moindre que la quantité de

travail maximale pouvant être produite. On peut alors se demander ce qui se passerait si ce processus était inversé. Selon la théorie de Carnot et la ST, un tel processus, qu'il soit cyclique ou non, procède tel une pompe à chaleur ou un réfrigérateur, c'est-à-dire qu'il consomme une quantité de travail afin de transférer de la chaleur du froid vers le chaud et de potentiellement intensifier une différence de température entre deux corps. Si, pour simplifier la discussion, l'on suppose que les corps (ou réservoirs) chaud et froid sont infinis, comme chez Carnot, alors, pour que le processus soit vraiment inversé, il faut que la quantité de travail consommée soit égale à celle produite dans le processus initial et que les quantités de chaleur transférées soient les mêmes, quoique dans un sens inverse. Et si, par hypothèse, le rendement du processus initial est moindre que le rendement maximal, il sera aussi moindre pour le processus inverse, de sorte que la quantité de travail sera insuffisante pour transférer les mêmes quantités de chaleur. En conséquence, les états initiaux ne sont pas récupérés et la théorie de Carnot et la ST, ainsi que les processus qu'elles décrivent, sont irréversibles au sens de l'irrécupérabilité.

Toutefois, il est possible en principe de récupérer les états initiaux, mais seulement au prix d'ouvrir le système et, donc, de modifier l'environnement (en consommant une quantité d'énergie supplémentaire). C'est-à-dire qu'il est possible d'utiliser des quantités de chaleur et de travail provenant d'« ailleurs » de manière à ce que l'état du système soit le même qu'initialement. En termes plus imagés, la « photo » du système, représentant son état, sera la même. Mais ce scénario ne répond pas aux critères de l'irréversibilité au sens de l'irrécupérabilité tel que définie précédemment. Comme discuté, ce qui permet d'établir les processus possibles en thermodynamique exige la définition et l'identification des systèmes en question et de leur environnement. Sans ces précisions à l'égard du système et de ses interactions avec son environnement, la question de savoir ce qui est possible pour un système devient très ambiguë. Cette ambiguïté est d'ailleurs à l'origine des difficultés d'interprétation quant aux systèmes ouverts où il est possible d'obtenir une diminution locale d'entropie, c'est-à-dire une diminution par rapport au système avec une augmentation d'entropie dans l'environnement : ce n'est parce que l'entropie d'un système donné diminue que nécessairement le second principe ou le principe entropique sont violés. C'est ce qui parut à prime abord étrange à ceux et celles qui ont tenté d'appliquer le second principe aux organismes vivants qui présentent une diminution *locale* d'entropie (section 4.2.5 et Chapitre 6).

Uffink (2001 : 323) arrive toutefois à une conclusion similaire, par rapport à l'irréversibilité décrite par la théorie de Carnot, par d'autres moyens, puisqu'il soutient qu'un « rendement moindre que le rendement maximal » signifie que le processus inverse agissant comme une pompe à chaleur demande *moins* de travail pour transférer les mêmes quantités de chaleur, et que par conséquent il serait possible de produire un mouvement perpétuel. Cette conséquence étant inadmissible, il en conclut que le processus inverse est impossible et, donc, que la théorie de Carnot est non-IRT. Mais si le rendement

est bien défini comme la quantité de travail obtenue d'une quantité de chaleur fournie au système, alors le rendement du processus inverse, c'est-à-dire en mode pompe à chaleur, doit être défini comme la quantité de chaleur transférée à partir d'une quantité de travail fournie au système. Un « rendement moindre que le rendement maximal » signifie donc que le processus produit moins de travail dans un sens, c'est-à-dire pour un transfert de chaleur positif où il y a plus de chaleur entrant dans le système que de chaleur y sortant, et qu'il en demande plus dans l'autre, c'est-à-dire pour un transfert de chaleur négatif où il y a moins de chaleur entrant dans le système que de chaleur y sortant. Il n'est donc pas possible d'obtenir un mouvement perpétuel avec deux processus dont le rendement est moindre que le rendement maximal et fonctionnant dans un sens inverse l'un de l'autre.

Pourtant, Uffink (2001 : 324-5) adresse plusieurs objections à la thèse voulant que la théorie de Carnot décrive des processus irréversibles au sens de l'irrécupérabilité. Et ces objections s'appliquent aussi dans une certaine mesure à la ST, d'où l'intérêt de les examiner. La première de ces objections concerne le cas limite que décrit la théorie où le processus présenterait un rendement maximal, ce qui correspond à un processus réversible chez Carnot et qui n'est pas spécifié dans la ST. En effet, dans ce cas limite d'un rendement maximal, la théorie de Carnot et la ST, tel que discuté, sont IRT. Or, comme le souligne Uffink, l'irréversibilité serait alors un ajout et non une conséquence de la théorie. Une comparaison avec la première loi de la mécanique newtonienne est à cet égard pertinente. Cette loi stipule qu'un corps libre, sans force s'y appliquant, persiste dans un état de mouvement rectiligne uniforme. Mais un corps libre est en ce sens, tout comme un processus réversible, seulement une idéalisation. Les corps « réels » sont toujours soumis à une force (ne serait-ce parce que la portée de l'attraction gravitationnelle est virtuellement infinie) et ils ne persistent donc pas dans un état de mouvement rectiligne uniforme. L'argument de Uffink veut que cette affirmation, à l'effet que les corps réels ne se déplacent pas selon un mouvement rectiligne uniforme ou seulement comme cas limite, soit un ajout à la première loi de la mécanique newtonienne et qu'il en serait de même de la théorie de Carnot et des processus irréversibles.

Sans entrer dans une discussion visant à déterminer s'il s'agit logiquement d'une conséquence ou d'un simple ajout, ce qui nécessiterait la présentation explicite de la théorie de manière axiomatisée, il semble bien qu'il soit de la nature même d'une description présentant un cas idéalisé de ne pas prétendre décrire littéralement des phénomènes « naturels » ou « réels »³². Mais, ce faisant, une telle description n'est pas nécessairement inopérante quant à l'explication de certains phénomènes

³² Il ne semble pas nécessaire ici de statuer sur un critère de réalité selon une position épistémologique particulière, autorisant ou non, par exemple, un discours sur les inobservables. Mais j'ajoute qu'un phénomène pouvant faire l'objet d'un test susceptible de corroborer ou de falsifier une théorie peut être considéré comme « réel », même s'il peut exemplifier une tendance décrite par une idéalisation exprimant un cas limite.

« naturels » ou « réels », comme les corps en mouvement soumis à la friction ou encore les échanges thermiques sans production de travail. Il y a plusieurs présentations en sciences de cas limite, de tendance asymptotique, où certaines propriétés des systèmes tendent dans certaines circonstances vers une certaine limite sans jamais l'atteindre et où cette tendance est justement décrite en fonction de cette limite. C'est le cas par exemple du postulat de la vitesse de la lumière, car les corps massiques ne peuvent atteindre cette vitesse limite à moins d'emmagasiner, par impossible, une énergie infinie, ou encore du zéro absolu, car un corps ne peut être refroidi à cette température à moins de subir, par impossible, une séquence infinie de cycles de refroidissement. Il semble bien que les instances exemplifiant ces comportements *approchant la limite* reçoivent une explication au sein des théories qui les décrivent et qui ont recours à ces idéalizations. Dans l'exemple précédent de la mécanique newtonienne, l'explanandum d'un corps qui n'est pas en mouvement rectiligne uniforme est jugé à l'aune de l'explanans de la première loi exprimant un cas limite³³. L'idéalisation qu'est ce cas limite permet de généraliser le concept de mouvement aux mouvements rectilignes et à ceux qui ne le sont pas (Einstein & Infeld 1963 : 22s). Elle paraît donc constitutive des descriptions des phénomènes que la théorie souhaite expliquer. Il en est de même pour les processus irréversibles : les processus peuvent tendre vers un rendement maximal mais ne l'atteignent jamais, car leur description ne répond pas à celle du cas limite, et lorsqu'ils ne répondent pas à cette description, c'est-à-dire *toujours*, alors ils sont irréversibles. En ce sens, la théorie de Canot de même que la ST peuvent rendre compte des phénomènes qui ne satisfont pas les critères du cas limite d'un processus réversible, et qui sont donc irréversibles.

Une autre objection veut que la théorie de Carnot ne présenterait pas de description de changements irréversibles dans l'état du système en question. L'objection semble être dirigée vers l'absence d'une fonction décrivant l'état du système, comme c'est le cas (ironiquement) avec la fonction d'état de l'entropie. Si c'est vraiment ce qui est recherché alors il est vrai que la théorie de Canot de même que la ST ne fournit pas une telle fonction. Mais il est étonnant de concentrer ainsi l'attention, et par conséquent l'objection, sur l'état des systèmes alors que la discussion était jusqu'à maintenant centrée sur les processus et les théories les décrivant. Il a été vu que la ST était réversible *qua* IRT seulement dans le cas limite jamais exemplifié d'un rendement maximal et irréversible *qua* irrécupérabilité si l'état de l'environnement était considéré, c'est-à-dire si la condition du système était imposée. Cela paraît tout à fait suffisant. Mais si des changements irréversibles dans l'état de l'environnement ou du système sont vraiment recherchés, alors une interprétation plus « indulgente »

³³ Un petit point technique ici : on pourrait objecter que les cas limite idéalisés n'expliquent pas mais permettent plutôt de « formuler » ou de « comprendre » l'explanandum. Ce genre d'objection n'est pas nécessairement invalide, mais il nécessite cependant une théorie précise et particulière de l'explication, ce qui ne semble pas nécessaire ici.

(« *liberal* », Uffink 2001 : 324) est possible. Il s'agit d'interpréter les flux de chaleur spontané entre deux corps comme une « perte de puissance motrice », une diminution de la capacité du système à effectuer un travail. Cette interprétation est applicable à un processus comme la ST puisqu'un processus inverse ne peut récupérer l'état initial. Que cette interprétation implique un énoncé « contrefactuel » (Uffink 2001 : 325), à savoir que si un processus présentant un rendement maximal avait été employé une quantité de travail maximale aurait été obtenue, ne pose pas de problème si l'on accepte, comme précédemment, que la description d'un cas limite, et donc idéalisé, puisse jouer un rôle majeur dans l'explication de certains phénomènes³⁴.

Alors, si la ST recouvre la théorie de Carnot et si elle décrit des processus irréversibles, peut-elle prétendre être l'élément constitutif du second principe de la thermodynamique ? Autrement dit, est-elle une candidate valable pour cerner la signification du second principe de la thermodynamique ? Une analyse des formulations de Clausius et de Kelvin (Chapitre 3) est éclairante à ce propos.

L'asymétrie dans les transferts de chaleur, discutée précédemment, est à l'origine de la distinction chez Clausius entre processus « spontané » et processus « forcé », ainsi que de sa formulation du second principe, généralement considérée comme son acte de naissance, laquelle se résume ainsi (Chapitre 3) : la chaleur ne peut passer d'elle-même d'un corps froid à un corps plus chaud ; et Clausius interprète « d'elle-même » par « sans compensation ». Ces deux expressions, au demeurant plutôt ambiguës, font référence au système et à son environnement, où l'expression « d'elle-même » réfère au système isolé et « sans compensation » correspond à d'éventuels changements dans l'environnement. En effet, tel que discutée, l'asymétrie se révèle en ce que la chaleur peut être transférée du chaud vers le froid « d'elle-même », sans intervention provenant de l'environnement, sans truchement d'un troisième système, tandis que le processus inverse, du froid vers le chaud, est impossible dans ces conditions. Plus précisément, ce transfert est seulement possible par une intervention provenant de l'environnement, par le truchement d'un troisième système prenant une quantité de chaleur du corps froid et la donnant au corps chaud. Encore une fois, sans définition du système et de ses éventuels interactions avec son environnement, la détermination de ce qui est possible ou non demeure ambiguë, voire impossible.

Cette intervention provenant de l'environnement, qui demande une dépense de travail sur le système, sous une forme ou une autre, car il y a plusieurs façons d'effectuer un travail, est aussi exprimée de plusieurs façons dans la littérature : « sans compensation » dans l'environnement (tel que

³⁴ Uffink (2001 : 325) présente aussi un argument voulant que si la théorie de Carnot implique l'irréversibilité, alors cela *doit* aussi être le cas des moulins à eau. Je ne vois cependant pas pourquoi l'*analogie* de Carnot, qui n'est donc pas une identité et qui se veut une illustration de sa théorie à partir d'un exemple connu (moulins), serait liée à une telle obligation. Ce n'est pas parce que des comportements sont similaires qu'ils doivent être identiques et ce n'est pas parce que des entités possèdent des comportements semblables qu'elles doivent posséder les mêmes propriétés. C'est justement le propre d'une analogie.

mentionné), « accompagnée d'un changement ailleurs », ou encore en affirmant l'impossibilité que ce processus de transfert de chaleur du froid au chaud soit « l'unique résultat » ou qu'il se fasse « de lui-même ». C'est pourquoi la formulation de Clausius a été reprise par plusieurs auteurs de manière semblable sans toutefois être identique (voir entre autres Van Wylen, Sonntag & Desrochers 1992 ; Lieb & Yngvason 2003 ; Atkins 2007). Relever les différences ou les équivalences exigerait une analyse minutieuse où seraient explicités les différentes prémisses sur lesquelles reposent ces différentes formulations. Mais, en termes simples, le second principe version Clausius dit que la chaleur ne peut être transférée dans la « mauvaise » direction, celle « forcée », que si un travail est fait. Il va sans dire que cet énoncé demande plus de précisions.

La ST stipule simplement qu'il y a transfert de chaleur du chaud vers le froid s'il y a une différence de température et qu'en certaines circonstances il peut y avoir production de travail. Inversement, elle stipule qu'il y a transfert de chaleur du froid vers le chaud s'il y a investissement de travail et qu'en certaines circonstances il peut y avoir une intensification de la différence de température. La clause « en certaines circonstances » signifie que les conditions de différence de température dans un cas et d'investissement de travail dans l'autre sont insuffisantes et qu'il faut donc quelque chose de plus. Il faut par exemple une machine à vapeur opérationnelle ou quelque chose du même genre pour que le flux de chaleur puisse produire un travail. La formulation par Clausius du second principe insiste ainsi sur une direction particulière des processus thermodynamiques, celui en mode réfrigérateur ou pompe à chaleur, autrement dit sur le mode « forcé » auquel réfère l'expression précédente de « mauvaise direction de la chaleur ». Le mode inverse, « spontané », n'exige en revanche qu'une différence de température et le transfert de chaleur s'effectue jusqu'à ce qu'il n'y ait plus de différence de température, autrement dit jusqu'à ce qu'il y ait équilibre thermique (bien que ce concept ne soit pas nécessaire pour l'instant). Or, la ST ne dit pas autre chose : le mode « forcé » exige une quantité de travail ($W \rightarrow Q \rightarrow \Delta T$), tandis que le mode « spontané » n'exige qu'une différence de température ($\Delta T \rightarrow Q$). Sur le strict point du processus de conversion de la chaleur en travail et du processus inverse, il y a donc convergence entre la formulation de Clausius et la ST.

La formulation de Kelvin insiste plutôt sur le cas limite du rendement maximal et se résume ainsi (Chapitre 3) : si une machine est telle que, lorsque son fonctionnement est renversé, ses éléments physiques et mécaniques sont tous renversés, alors elle produit autant de travail qu'il est possible d'en produire à partir d'une quantité de chaleur donnée. Cette formulation affirme (i) que les machines réversibles possèdent un rendement maximal et produisent donc une quantité de travail maximale pour une quantité de chaleur donnée, et (ii) que le rendement ne dépend que des températures d'opération. À l'instar de la formulation de Clausius, elle a été discutée et interprétée à maintes reprises (voir entre autres Van Wylen, Sonntag & Desrochers 1992 ; MacDonald 1995 ; Lieb & Yngvason 2003 ; Capek &

Sheehan 2005 ; Atkins 2007 ; Kondepudi 2008). En termes simples, le second principe version Kelvin dit qu'une quantité de chaleur ne peut être complètement convertie en travail. De fait, différentes versions de cet énoncé font valoir entre autres l'impossibilité d'une production de travail « d'une quantité égale » à celle de la chaleur ou à partir d'un échange de chaleur « avec un seul réservoir ».

Étant donné la formulation de Carnot à propos du rendement maximal possible, cette formulation par Kelvin du second principe implique aussi que ce maximum de travail pouvant être obtenue d'un système est strictement inférieur à la quantité de chaleur transférée à ce même système. Mais comme la chaleur ne peut être annihilée, une partie de cette chaleur est nécessairement rejetée dans l'environnement. C'est pourquoi, de manière imagée, on peut dire que la nature impose une « taxe » sur la conversion de la chaleur en travail, puisque, de la chaleur transférée au système, une partie de cette chaleur est nécessairement rejetée dans l'environnement, lequel est un élément nécessaire à la production de travail³⁵. Le premier principe établit seulement que cette quantité de travail maximale ne peut être supérieure à la quantité de chaleur transférée à ce système. Cependant, en l'état de l'énoncé, rien n'est dit de la provenance ou de l'origine de cette quantité de chaleur, car elle pourrait provenir d'un corps plus froid ou d'un corps plus chaud que le système en question. C'est pourquoi Kelvin ajoute son axiome à l'effet que la quantité de travail ne peut être obtenue à partir d'un corps qui serait refroidi à une température inférieure à celle des objets environnants. En effet, cela impliquerait qu'il y ait transfert de chaleur jusqu'à ce que la différence de température soit nulle, autrement dit jusqu'à l'équilibre thermique, puis de nouveau dans le sens inverse augmentant ainsi la différence de température entre le corps chaud et le corps froid. Or la ST interdit clairement une telle situation. Non seulement il ne peut y avoir de transfert thermique lorsque la différence de température entre deux corps est nulle, mais il ne peut y avoir de transfert direct de chaleur du froid vers le chaud, ni de transfert indirect de chaleur du froid vers le chaud, c'est-à-dire par le truchement d'un troisième corps, sans dépense de travail. Sur ce strict point du processus de conversion maximale de la chaleur en travail, donc, il y a convergence entre la formulation de Kelvin et la ST.

³⁵ Cet aspect péjoratif exprimé par le mot « taxe », qui se veut davantage illustratif que démonstratif, est relatif pour deux raisons. D'une part, cette chaleur « perdue » peut bien être utilisée pour chauffer un corps quelconque et elle n'est donc pas vraiment perdue, et ne représente donc pas vraiment une « taxe » si elle trouve une « utilité ». D'autre part, cette perte est relative à l'équivalence établie entre la chaleur et le travail, car sans cette équivalence, cette perte ne ferait aucun sens et on dirait seulement que de la chaleur est convertie en travail. C'est uniquement en raison de cette équivalence numérique et comptable dans certaines circonstances établies expérimentalement, où il peut y avoir une conversion complète du travail en chaleur, que la notion de perte de chaleur a un sens lorsqu'il y a une conversion inverse incomplète de la chaleur en travail. Cependant, il y a un autre sens rattaché à la notion de « perte » mais cette fois par rapport à une quantité de travail que l'on pourrait qualifier de « contrefactuelle » lorsque celle-ci est inférieure à la quantité maximale qui aurait pu être produite mais qui ne l'a pas été (discuté plus loin).

L'argumentation précédente permet donc de présenter la séquence thermodynamique (ST) comme une candidate valable pour cerner la signification du second principe de la thermodynamique. Qu'elle puisse vraiment en dégager l'« essence » ou non demeure une question plus difficile. Il est tout de même raisonnable d'affirmer qu'elle limite les hypothèses et qu'elle offre une synthèse adéquate des énoncés de Kelvin et Clausius.

4.2.3.3 CRITIQUES

Un reproche maintes fois adressé contre les versions du second principe de Clausius et de Kelvin concerne le caractère « négatif » des énoncés, en ce sens où ils formuleraient ce qui est impossible plutôt que ce qui est possible (par ex. Van Wylen, Sonntag & Desrochers 1992 ; Çengel & Boles 2006). De fait, personne n'a jamais observé un phénomène entrant en contradiction avec le second principe de la thermodynamique ; plus spécifiquement, personne n'a jamais observé un corps se réchauffer de lui-même, ni une quantité de chaleur être complètement convertie en travail. La thermodynamique ne traite évidemment pas des « miracles », dont le sens est intimement lié à la notion de « possible » (discutée précédemment ; voir aussi Earman 2000, Uffink 2001 et McGrew 2011). Si tant est que le second principe puisse être une « loi de la nature », il est tout à fait trivial qu'il interdise les miracles entendus comme « violation des lois de la nature » selon la formule consacrée de Hume (1748). En ce sens, la prétention du second principe à décrire les phénomènes naturels est aussi une prétention contre la réalité des miracles. De manière plus restreinte, un miracle peut être entendu comme ce qui est au-delà de la portée de l'action humaine, bref de ce qui nous est impossible d'accomplir, et en ce sens le second principe exprimerait nos propres limitations par ce caractère négatif. Il est aussi possible d'interpréter la notion de miracle comme quelque chose d'« inexplicable par des causes naturelles ». Mais, dans tous les cas, une observation ou une preuve théorique ou tout autre moyen jugé épistémologiquement fondé, et qui serait en contradiction avec l'énoncé du second principe de la thermodynamique, constituerait un candidat sérieux pour invalider la prétention de ce principe à rendre compte des phénomènes naturels dont traitent les sciences. En attendant, cette prétention reste intacte.

Une conséquence de ce caractère négatif serait l'impossibilité d'exprimer une tendance à l'irréversibilité des phénomènes (Brush 1976 : 571). Le problème semble cosmétique puisqu'il est tout à fait possible de reformuler ces énoncés de manière « positive » par un choix approprié de vocabulaire. D'autant plus qu'il existe plusieurs principes scientifiques aussi formulés en termes « négatifs », que ce soit le postulat de la vitesse de la lumière, le principe d'indétermination quantique, ou même le troisième principe de la thermodynamique. Une autre conséquence serait l'impossibilité de présenter une « preuve » de sa validité. Cela signifie en fait qu'il n'y aurait aucune corroboration expérimentale

possible d'un énoncé statuant sur ce qui est impossible. Il est généralement admis, toutefois, que ce qui importe n'est pas tant la corroboration d'un énoncé scientifique que sa falsification, et plus précisément l'absence de falsification. Un énoncé serait ainsi considéré comme vrai jusqu'à preuve du contraire. Mais si cette preuve est par principe impossible, si l'énoncé est compatible avec toute observation possible, alors l'énoncé est jugé infalsifiable et (donc) non scientifique. Ici, il n'y a pas d'impossibilité principielle d'une telle preuve, c'est-à-dire d'une preuve d'un phénomène dont la description entrerait en contradiction avec l'énoncé du second principe, seulement une impossibilité principielle, ou plutôt formelle par ce caractère négatif, d'une corroboration expérimentale. Or il n'y a pas de danger de qualifier les énoncés du second principe de Clausius et de Kelvin, ni la ST, de non falsifiables car ils ne sont pas compatibles avec toute observation possible. Cela étant dit, jusqu'à maintenant, aucune observation n'a mené à une contradiction avec ni l'énoncé de Clausius ni celui de Kelvin.

La formulation du second principe en termes « négatifs » n'est donc pas un problème quant à ses prétentions d'expliquer certains phénomènes thermodynamiques, à moins d'avoir en main justement une théorie de l'explication toute particulière et novatrice interdisant une formulation « négative ». Mais hormis ce caractère négatif, une autre caractérisation ayant trait à la notion de « possible » est nécessaire. Plus spécifiquement, il s'agit de l'expression d'un *conditionnel contrefactuel*, où typiquement quelque chose serait ou aurait été possible, ou se serait produit, si les circonstances étaient ou avait été différentes³⁶. En exprimant un contrefactuel, le second principe décrirait ce qui se produirait dans un monde possible, et non dans ce monde, ce qu'exprimerait un énoncé factuel. La description d'un processus présentant un rendement moindre que le rendement maximal exprime ainsi le contrefactuel à l'effet qu'il aurait été possible d'extraire une quantité supérieure de travail si le rendement avait été supérieur, si la quantité de chaleur cédée à l'environnement n'avait pas été si élevée. De manière plus imagée, et pour reprendre les termes précédents, les choses auraient été différentes si la « taxe », ce « tribut » payé à l'environnement pour la conversion de la chaleur en travail, n'était pas si élevée. En l'occurrence, cela correspond à une « perte » de capacité à effectuer un travail. Cependant, ici, le passage d'un tel énoncé contrefactuel à un énoncé ayant une prétention factuelle est impossible pour la simple raison qu'une telle situation de rendement maximal est impossible. Selon les distinctions apportées précédemment, cela est théoriquement, réellement (dans les faits) et humainement impossible. En effet, il s'agit d'un cas limite, une idéalisation jamais exemplifiée. Or, la capacité à supporter des énoncés contrefactuels est généralement considérée comme un trait particulier des lois (Hempel 1965 ; Horwich 1987) ou des relations causales³⁷ (Woodward 2003 ; Psillos 2007). Ce n'est

³⁶ Un conditionnel contrefactuel est un conditionnel subjonctif avec un antécédent faux.

³⁷ Comme quoi l'interprétation causale de la ST serait justifiée.

donc pas un problème pour le second principe de la thermodynamique selon la formulation de la ST quant à sa prétention à expliquer certains phénomènes. Le problème devrait donc être ailleurs.

Un autre problème associé à la formulation contrefactuelle concernerait l'argument à l'effet que le second principe décrive des processus irréversibles au sens de changements irrécupérables. Car l'énoncé conditionnel contrefactuel exprime une possibilité et non un fait, c'est-à-dire que, si les conditions étaient ou avaient été différentes, alors il serait ou aurait été *possible* de produire un certain état de faits, et non pas que tel ou tel événement se serait effectivement produit. Autrement dit, l'antécédent *et* le conséquent du conditionnel expriment une possibilité. Ce contrefactuel ne concerne donc pas un monde où le second principe n'existerait pas et où il serait possible de convertir complètement une quantité de chaleur en travail, comme le laisse entendre le premier principe, mais plutôt un monde où l'application du principe aurait été différente. Par exemple, si un transfert de chaleur s'effectue « spontanément » d'un corps chaud à un corps froid, donc sans le truchement d'un autre corps, il a été vu que ce processus est irréversible au sens de non-IRT et d'irrécupérabilité pour autant que le système soit bien défini. Mais il aurait été possible d'exploiter la différence de température entre ces deux corps afin de produire un travail. Et dès que le transfert s'opère la différence de température diminue et la capacité d'effectuer un travail aussi. Il aurait donc été possible d'avoir un certain état de faits, c'est-à-dire une quantité de travail, si ce transfert de chaleur avait été exploité. En ce sens, ce phénomène est similaire à la perte d'un potentiel ; et cela est naturel pour des quantités n'étant pas représentées par une fonction d'état, comme le sont la chaleur et le travail, c'est-à-dire des quantités qui dépendent du chemin parcouru, du processus impliqué. Mais, dès que ce transfert est effectué, il est « trop tard », le « mal est fait », le phénomène est irréversible selon le sens donné précédemment, soit une irrécupérabilité, à moins d'ouvrir les limites du système. Voilà le conditionnel contrefactuel qu'exprime le second principe. Il faut donc comprendre qu'il y a bien un changement irrécupérable, pour autant que le système soit bien défini.

Cette conclusion a des conséquences pratiques importantes car elle exprime ce « fait physique » de l'irrécupérabilité de la capacité de travail d'un système, laquelle ne peut être retrouvée qu'en ouvrant les limites de ce système, bref en consommant des ressources à l'extérieure de celui-ci. Si la capacité d'un système à effectuer un travail possède une valeur monnayable dans une société industrielle comme la nôtre, alors il appert que la consommation de ressources pour maintenir cette capacité se produira d'une façon ou d'une autre³⁸.

³⁸ Ce point est évidemment important par rapport à l'écologie et à la notion de développement durable. Sans entrer dans une discussion sur la signification de cette notion largement galvaudée, l'idée ici est la suivante : la société industrielle a besoin de sources d'énergie convertible sous forme de travail, donc à faible entropie, et si la consommation d'énergie entraîne une

Il y a, par ailleurs, un autre conditionnel contrefactuel exprimé par le second principe. Il ne concerne pas la possibilité d'obtenir un certain état de faits à partir de ce système si une certaine intervention sur celui-ci avait été menée, comme dans le cas précédent. Il concerne plutôt la possibilité d'obtenir un certain état de faits à partir de ce système si l'environnement avait été différent. La nuance peut sembler minime, mais elle demeure importante face à la question de savoir ce qui est possible pour un système thermodynamique en fonction de son environnement. Puisqu'un système est toujours dans un certain environnement et que celui-ci détermine dans une large mesure ce que le système peut accomplir, une caractérisation en faisant abstraction semble une stratégie mal avisée. Plus tôt, le contrefactuel exprimait la possibilité d'une certaine quantité de travail maximale pour un état donné du système, mais il est aussi évident qu'il est possible d'en exploiter moins que cette quantité maximale. Ici, le contrefactuel exprime plutôt la possibilité d'une certaine quantité de travail en fonction de l'environnement. Par exemple, si le système est caractérisé par une certaine température, la quantité de travail maximale qui est possible d'en tirer dépend de la température de l'environnement, car plus la différence de température sera grande, plus grande sera la quantité de travail maximale pouvant être obtenue de ce complexe système-environnement. Cela signifie que la caractérisation de l'état du système quant à ses capacités thermodynamiques sont relatives à l'environnement, à l'instar de l'adaptabilité (« *fitness* ») dont la valeur est aussi relative à un environnement.

Un dernier point avant de conclure cette section : les formulations de Clausius et Kelvin sont-elles équivalentes ? L'équivalence entre ces deux énoncés a été maintes fois alléguée (par ex. Albert 2000 ; Çengel & Boles 2006 ; Atkins 2007). Un critère jugé suffisant pour que ces deux formulations soient équivalentes est qu'elles doivent être vraies en même temps et fausses en même temps, c'est-à-dire dans les mêmes circonstances. En d'autres mots, la violation d'un principe impliquerait la violation de l'autre. L'argument repose généralement sur le couplage, jugé anodin, d'une machine violant l'une des formulations du second principe, une machine « anti-Clausius » ou « anti-Kelvin » selon le cas, avec une machine « normale » fonctionnant selon un cycle de Carnot. Le fonctionnement de ces machines combinées mènerait à la violation de l'autre version du principe, celui ne faisant pas l'objet de la machine impossible dans l'exemple. Une objection à cet argument consiste à dire que l'ordonnancement des températures influence la validité de l'argument, car si on fait usage de températures négatives³⁹, l'argument ne tiendrait plus (Ehrenfest-Afanassjewa 1925, in Uffink 2001 ;

augmentation importante d'entropie, alors le recouvrement de cette capacité de travail implique d'ouvrir le système et donc de consommer d'avantage de ressources ; on peut même pousser le raisonnement jusqu'à identifier l'entropie avec la pollution, sous forme de déchets inutilisables ou encore sous forme de chaleur. (Ce point sera discuté au Chapitre 6 ; voir aussi Fleissner & Hofkirchner 1997 ; Hammond 2004 ; Vivien 2005 ; Deutscher 2008 ; Kümmel 2011.)

³⁹ Voilà une notion pour le moins hétérodoxe puisque selon le troisième principe il est impossible d'obtenir une température inférieure au zéro absolu (qui porte bien son nom). On aurait pu sans douter, le signe négatif apparaît en raison de certaines

Ramsey 1956). Mais, encore une fois, le problème n'est qu'apparent puisque l'ordonnancement du chaud et du froid peut être maintenu pour les températures négatives avec un choix approprié de vocabulaire. La validité du couplage, en revanche, est plus problématique, car on ne peut prétendre qu'elle aille de soi. Mais si l'on peut prouver l'équivalence de ces énoncés avec un autre énoncé jugé valable, alors le résultat souhaité serait obtenu. Et si l'argumentation précédente à propos de la ST est correcte, alors ce résultat quant à l'équivalence est disponible.

En conclusion. Le second principe de la thermodynamique se propose de mettre un terme à ce « désir pathétique » d'obtenir quelque chose à partir de rien. Plus précisément, il interdit une conversion complète de la chaleur en travail. En ce sens, il décrit l'impossibilité d'un mouvement perpétuel de second type, soit un processus permettant de convertir toute la chaleur d'un corps en travail. La robustesse de ce principe exprimant cette impossibilité n'est plus remise en doute et elle résulte principalement de raisons de fait plutôt que de principe. La signification du second principe a été dégagée ici par une reconstruction théorique basée sur quelques principes et appelée *séquence thermodynamique* (ST), qui stipule essentiellement qu'une quantité de chaleur ne peut être complètement convertie en travail. Par conséquent, le second principe n'est pertinent qu'à conditions que les concepts de chaleur et de travail soient bien définis. Bien qu'il soit généralement admis que ces derniers définissent les deux modes généraux de transferts d'énergie, il est loin d'être évident que le second principe puisse s'appliquer tel quel aux phénomènes microscopiques. Une généralisation à « tous les processus de la nature » ne va donc pas de soi.

Avec cette tendance à la diminution de la capacité à effectuer un travail et à l'uniformisation des températures, il y a une directionnalité aux processus énergétique appliqués aux systèmes thermodynamiques puisque l'inverse est possible, c'est-à-dire qu'une quantité de travail peut être complètement convertie en chaleur. Il a été vu que le caractère « négatif » de ce type d'énoncé ainsi que l'expression d'un contrefactuel ne posaient pas de problème majeur. En effet, le second principe énonce un cas limite jamais instancié et dans les faits cette limite n'est jamais atteinte. De plus, il adopte une position épistémologique très similaire à d'autres principes scientifiques comme la première loi de la mécanique newtonienne. Enfin, il n'est pas infalsifiable car il suffirait d'observer un phénomène dont la description correspondrait à celle du cas limite idéalisé, puisqu'un tel cas est, *ex hypothesi*, impossible.

caractéristiques quantiques particulières. Il faut alors comprendre que la température est un paramètre exprimant la population relative d'occupation des niveaux d'énergie quantique dans un système à l'équilibre. Normalement, c'est-à-dire pour les températures positives, les niveaux de basse énergie sont plus occupés que les niveaux de haute énergie. Or, pour les températures négatives, c'est l'inverse. Ces valeurs négatives ne signifient toutefois pas que ces températures sont plus « froides » que leurs contreparties positives. Par contre, dans ces situations particulières, l'entropie *diminue* lorsqu'on ajoute de l'énergie. De plus, on ne peut en conclure que le système est passé par le zéro absolu, qui reste, donc, impossible à atteindre.

Les critiques ne portent donc pas tant sur le principe lui-même que sur cette stratégie explicative basée sur un cas limite idéalisé.

Cette directionnalité exprime une asymétrie temporelle selon les deux sens donnés au concept d'irréversibilité. Ainsi, un processus thermodynamique est non-invariant par renversement du temps (non-IRT) car aucun processus ne peut satisfaire ce cas limite d'un rendement maximal où il y aurait interconvertibilité complète de la chaleur et du travail à l'intérieur d'un système donné. Un processus thermodynamique est aussi irréversible au sens d'irré récupérable car l'unique façon de récupérer un état initial est d'ouvrir le système et d'échanger de l'énergie, et éventuellement de la matière avec l'environnement. Cela revient à modifier le système et l'environnement, ce qui ne peut satisfaire le critère d'irréversibilité au sens d'irré récupérabilité préalablement défini.

4.2.4 La fonction d'état S

Le besoin d'exprimer le second principe de la thermodynamique par un énoncé quantitatif et mathématisé est sans doute révélateur de son importance. Non seulement ce besoin se confond avec un désir depuis la révolution scientifique, le paradigme de la théorie newtonienne axiomatisée, hypothético-déductive, s'étant imposé, mais aussi la généralité qu'offre une formulation mathématique permet d'unifier jusqu'à un certain point la diversité rencontrée dans les énoncés du second principe. Mais puisque ce principe est un condensé d'observations empiriques, cette formulation mathématique demande des hypothèses supplémentaires qui ne retrouvent pas nécessairement la robustesse offerte par la sanction de l'expérience dont ce principe se veut justement le résumé. C'est pourquoi on peut bien soutenir que les gains recherchés ne balancent pas nécessairement ce qui a été perdu. La discussion qui suit présente ces hypothèses à la base de la définition de l'entropie et les propriétés qui en découlent.

CONSTRUCTION

Les étapes de l'argumentation menant à la définition de l'entropie sont très rarement présentées dans la littérature (une exception notable : Fermi 1936). Mais si l'utilité du résultat peut justifier jusqu'à un certain point sa validité, la façon d'y arriver a bien sûr son importance. Ainsi la présentation de ces étapes permet une meilleure compréhension de ce même résultat, tant en ce qui a trait à sa signification qu'à sa portée.

1. Clausius, à l'instar de Carnot, suppose l'existence d'une fonction des températures d'opération (T et T_+) déterminant la quantité de travail maximale pouvant être produite d'une quantité de chaleur

donnée. Il considère d'abord le cas plus simple des processus cycliques fermés où l'état initial est récupéré, car certains changements subis lors de ce processus peuvent alors s'annuler. L'argument est alors qu'il existe des « valeurs d'équivalence » représentant mathématiquement ces changements subis lors de ce processus cyclique et que celles-ci s'annulent à la fin de ce processus. Son raisonnement s'appuie ensuite sur le principe et le corollaire de Carnot, comme quoi cette quantité de travail maximale obtenue d'une quantité de chaleur donnée est indépendante de la substance et dépend de la température. Les « valeurs d'équivalence » prennent ainsi la forme suivante : $f(T_i)Q_i$, où $f(T_i)$ est une fonction universelle indéterminée de la température. La condition de réversibilité *qua* récupérabilité demande que ces valeurs se compensent mutuellement pour un cycle complet, de sorte que

$$\sum f(T_i)Q_i = 0 \text{ [processus cyclique réversible].}$$

Lorsque le système suit un processus impliquant des quantités infinitésimales de chaleur δQ , c'est-à-dire lorsqu'il subit un cycle quasi-statique dans lequel les réservoirs de chaleur ont une température variant continument, autrement dit lorsque le nombre d'éléments de la somme dans l'équation précédente tend vers l'infini, alors cette somme peut être exprimée par une intégrale

$$\oint f(T) \delta Q = 0 \text{ [processus cyclique réversible].}$$

Notons que l'intégrale \oint et la différentielle δ , en notation moderne, représentent respectivement une intégrale cyclique et une différentielle inexacte⁴⁰. Notons par ailleurs que « isolé » signifie plus précisément « adiabatiquement isolé », c'est-à-dire sans échange d'énergie sous forme de chaleur, et que T représente la température absolue de transfert de la chaleur, du corps recevant la chaleur.

En considérant que l'intégrale précédente est indépendante du chemin d'intégration, c'est-à-dire qu'elle dépend uniquement des états initial x_i et final x_f , il est possible (comme l'a fait Clausius) d'établir mathématiquement l'existence d'une *fonction d'état* S . En effet, si cette fonction est appliquée à un processus cyclique réversible et adiabatiquement isolé, d'un état initial x_i à un état final x_f par un processus puis de retour à x_i par un autre processus, alors cette intégrale peut être séparée en deux composantes dont la somme est nulle. La fonction d'état est donc

$$\oint f(T)\delta Q = \int dS = S(x_f) - S(x_i) = 0 \text{ [processus cyclique réversible].}$$

Il reste donc à déterminer la fonction $f(T)$ correspondant à la fonction de Carnot (f_C). La détermination de cette fonction s'appuie généralement sur la construction d'un modèle faisant intervenir deux machines opérant un cycle de Carnot et appliqué à un gaz parfait, donc sur deux

⁴⁰ Clausius n'utilisait pas de notation spéciale pour ces termes.

idéalisations, bien qu'il soit possible de le généraliser au-delà du domaine des gaz parfaits. Ce modèle s'appuie sur une version du second principe, comme quoi un transfert de chaleur du froid vers le chaud sans être accompagné d'une consommation de travail est impossible. Or, le rendement (η) des machines est défini comme le rapport de la quantité de travail (W) produite sur la quantité de chaleur consommée (Q_+). Et selon le théorème de Carnot, ce rendement ne dépend que des températures d'opération du cycle ($\eta = f_C(T, T_+)$). Une étape dans l'argumentation indique que la fonction de Carnot d'une machine fonctionnant entre deux températures (T_- et T_+) peut s'exprimer sous la forme d'un ratio tel que $f_C(T, T_+) = f(T_-)/f(T_+)$. Une autre étape consiste à définir la fonction $f(T)$ comme la température absolue selon l'échelle de Kelvin. De sorte $f(T_-)/f(T_+) = T_-/T_+$, et, donc, $\eta = 1 - T_-/T_+$. On peut alors montrer facilement que $\sum f(T_i)Q_i = \sum Q_i/T_i = 0$, où T_i correspond à la température absolue. (Pour plus de détails sur ce type d'argument, voir Truesdell 1980 ; Jaynes 1988 ; Uffink 2001 ; Müller 2007 ; Kondepudi 2008 ; Müller & Müller 2009.) Avec ce résultat, on obtient

$$\int dS = \int \delta Q/T \text{ [processus réversible]}.$$

2. Le résultat précédent de Clausius peut être reformulé de la manière suivante : la différentielle inexacte (incomplète, non-intégrable) δQ possède un diviseur intégrant (T) qui la transforme en différentielle exacte (complète, intégrable) et donc il existe une fonction d'état telle que

$$\int \delta Q/T \equiv \int dS = S(x_i) - S(x_f) \text{ [processus réversible]}.$$

Puisqu'il s'agit d'une fonction d'état, pour un processus cyclique réversible il s'ensuit que la valeur de la fonction d'état à l'état initial et la même qu'à l'état finale, donc $S(x_i) = S(x_f)$ et la valeur de l'intégrale définie précédente est nulle. Autrement dit, l'intégrale de $\delta Q/T$ sur un chemin reliant deux états est ainsi indépendante du chemin emprunté et dépend seulement de ces états, qui doivent être suffisamment près de l'équilibre. En effet, pour un processus cyclique réversible, l'état initial du système est par définition le même que son état final et par conséquent l'entropie initiale est égale à l'entropie finale de ce système. Il est alors possible de démontrer que la valeur de la quantité $\int \delta Q/T$ pour deux états à l'équilibre est la même peu importe le processus reliant ces deux états⁴¹.

3. Il s'agit ensuite de déterminer cette fonction d'état pour des processus, cycliques ou non, irréversibles. L'argument de Clausius est à peu près le suivant : deux machines effectuant un cycle sont couplées, l'une selon un cycle réversible et l'autre selon un cycle irréversible ; alors la machine

⁴¹ Soit deux états à l'équilibre x_i et x_f et deux processus réversibles, A allant de x_i et x_f , et B allant de x_f et x_i ; alors il s'agit d'un cycle et par conséquent la quantité $\int \delta Q/T$ est nulle et, donc, l'entropie de l'état initial est égal à l'entropie de l'état final. Plus précisément, le processus cyclique implique que $\int [A : x_i \rightarrow x_f, B : x_f \rightarrow x_i] \delta Q/T = 0$; ce qui revient à scinder l'intégrale $\int [A : x_i \rightarrow x_f] \delta Q/T + \int [B : x_f \rightarrow x_i] \delta Q/T = 0$; donc $\int [A : x_i \rightarrow x_f] \delta Q/T = -\int [B : x_f \rightarrow x_i] \delta Q/T = \int [B : x_f \rightarrow x_i] \delta Q/T$.

fonctionnant selon un cycle irréversible ne peut compenser les transformations opérées par la machine fonctionnant selon un cycle réversible, c'est-à-dire qu'elle ne peut fournir suffisamment de travail ou tirer suffisamment de chaleur, selon le processus considéré ; de sorte $\oint dQ/T \neq 0$; et par la convention de signe $\oint dQ/T < 0$. L'idée générale est donc qu'une certaine quantité, déterminée par le passage d'un système d'un état initial à un état final par un certain processus, ne puisse être compensée par une autre quantité, déterminée cette fois par le passage de ce système de l'état final précédent à l'état initial original par un processus différent. Un processus cyclique est alors défini par une combinaison d'un processus réversible et d'un processus irréversible. L'ajout d'un processus irréversible est ainsi une condition rédhibitoire à la réversibilité de l'ensemble, en ce sens où un seul processus irréversible dans une suite de processus réversibles suffit à considérer l'ensemble total des processus comme étant irréversible. Ainsi, si l'on considère un processus cyclique irréversible, tel que

$$\{\oint [x_i \rightarrow x_f] \delta Q/T\}_{\text{irrév}} + \{\oint [x_f \rightarrow x_i] \delta Q/T\}_{\text{rév}} < 0,$$

et puisque $\{\oint [x_f \rightarrow x_i] \delta Q/T\}_{\text{rév}} = S(x_i) - S(x_f)$, on a alors $\{\oint [x_i \rightarrow x_f] \delta Q/T\}_{\text{irrév}} + S(x_i) - S(x_f) < 0$. Donc,

$$\{\oint [x_i \rightarrow x_f] \delta Q/T\}_{\text{irrév}} < S(x_f) - S(x_i).$$

Ainsi,

$$\oint \delta Q/T < S(x_f) - S(x_i) \text{ [processus irréversible]}.$$

Cela signifie que pour tout processus se produisant dans un système adiabatiquement isolé, l'entropie de l'état final $S(x_f)$ est plus grande ou égale à l'entropie de l'état initial $S(x_i)$. En somme, l'entropie thermodynamique, S_T ⁴², est définie telle que

$$S_T(x_f) - S_T(x_i) = \oint \delta Q/T \quad [\text{si } x_i \rightarrow x_f \text{ est réversible}],$$

$$S_T(x_f) - S_T(x_i) > \oint \delta Q/T \quad [\text{si } x_i \rightarrow x_f \text{ est irréversible}].$$

Un élément important dans cette argumentation est l'existence d'un processus réversible, $\{\oint [x_f \rightarrow x_i] \delta Q/T\}_{\text{rév}}$, pouvant boucler le cycle irréversible afin que la quantité $\{S(x_i) - S(x_f)\}_{\text{rév}}$ soit bien définie. Mais il n'est pas évident que l'hypothèse de l'existence d'un tel processus puisse être généralisée à des systèmes plus complexes que ceux définis dans l'argumentation, à savoir des systèmes respectant la description d'un gaz parfait. C'est pourquoi la généralisation de Clausius (1865, in 1868 : 419) à « toutes les transformations qui s'effectuent dans la nature » paraît « téméraire » (Uffink 2001 : 339).

Cependant, la témérité de Clausius est davantage liée à la définition de sa fonction d'état qu'à la prétendue existence d'un processus réversible. En effet, s'il est entendu que les processus irréversibles

⁴² Dans un contexte où il n'y a pas d'ambiguïté quant à la définition de l'entropie, le symbole S sera simplement utilisé.

présentent des rendements inférieurs aux processus réversibles, alors il s'ensuit que ces processus irréversibles présente un rendement ($\eta_{\text{irr}}^{\text{év}}$) moindre que celui défini par la formule de Carnot ($\eta_{\text{rév}}^{\text{év}}$) pour des températures données. C'est-à-dire que, par les définitions même de ces rendements, $\eta_{\text{irr}}^{\text{év}} = 1 - Q_-/Q_+ < \eta_{\text{rév}}^{\text{év}} = 1 - T_-/T_+$; ce qui implique que $Q_-/Q_+ > T_-/T_+$, donc $Q_+/T_+ < Q_-/T_-$ pour un processus irréversible ayant un rendement moindre que le rendement maximal défini pour un processus réversible. En clair, si la fonction d'état censée représenter cette nouvelle variable qu'est l'entropie est valide et qu'elle s'applique au système en question, alors cette dernière inégalité indique que lors d'un processus irréversible l'entropie cédée au corps froid (Q_-/T_-) est supérieure à celle cédée par le corps chaud (Q_+/T_+) et que, par conséquent, l'entropie globale augmente. C'est-à-dire que, plus précisément, elle ne peut être nulle et par convention de signes elle augmente. L'égalité $Q_+/T_+ = Q_-/T_-$ est respectée que dans le cas limite d'un processus réversible. Mais ce cas limite n'est, par définition, et l'expérience l'atteste, jamais exemplifié. Ce n'est donc pas l'existence d'un processus réversible qui est en jeu, mais plutôt l'écart qui sépare tout processus de ce cas limite définissant la fonction d'état (section 4.2.7).

Il est évident que l'on peut remettre en cause l'argumentation faisant état de l'irréversibilité de la théorie, d'après, bien entendu, une interprétation claire de ce qui est signifié par ce terme. Mais la signification de l'irréversibilité est déterminée ici par la théorie qui définit les conditions d'un processus réversible. Si un processus ne remplit pas ces conditions, alors, trivialement, il n'est pas réversible. Et si un processus ne remplit pas ces conditions, alors il possède certaines caractéristiques déterminées par la théorie; par exemple, $Q_+/T_+ > Q_-/T_-$. Or, encore une fois, la théorie définit les conditions d'un processus réversible comme un cas limite jamais instancié. Ce sont donc les processus qui ne sont pas réversibles qui sont susceptibles d'être observés et leur description corroborée. Une éventuelle critique devrait ainsi porter non pas sur la signification de l'irréversibilité mais plutôt sur cette stratégie explicative visant à définir un cas limite idéalisé par rapport auquel les processus susceptibles de faire l'objet d'une corroboration sont évalués. C'est en ce sens seulement que les processus qui ne sont pas réversibles, donc irréversibles, sont « réels ». Sans cet ingrédient qu'est cette séparation entre un cas limite idéalisé et les autres cas qui sont évalués par rapport à celui-ci, la critique a tout le loisir d'affirmer que la théorie ne traite pas d'irréversibilité.

4. Lieb & Yngvason (1999 ; 2003) ont aussi défini une fonction d'état mais à partir d'une quinzaine d'axiomes. L'ingrédient principal de leur reconstruction conceptuelle est une certaine relation d'ordre définie sur un espace des phases, appelée *accessibilité adiabatique*, dénotée par $<$ et prononcée « précède ». Un état T est adiabatiquement accessible d'un état X s'il est possible de passer d'un état X à un état T au moyen d'une interaction avec un certain dispositif consistant en un système auxiliaire et un poids, de telle façon à ce que le système auxiliaire retourne à son état initial à la fin du processus, alors que le poids peut avoir monté ou descendu dans un champ gravitationnel. Par exemple, un système dans un

état X peut être composé d'un verre d'eau chaude et d'un verre d'eau glacée ; il est alors possible de relier ces verres d'eau par un système auxiliaire comme un fil conducteur et de passer à un état \mathcal{I} où les deux verres sont dans un équilibre thermique, c'est-à-dire à la même température, alors que le fil conducteur est retourné à son état initial ; par conséquent \mathcal{I} est « adiabaticquement accessible » à partir de X : $X < \mathcal{I}$. Par contre, dans cet exemple, ce système auxiliaire serait insuffisant pour passer de \mathcal{I} à X ; par conséquent X n'est pas « adiabaticquement accessible » à partir de \mathcal{I} : $\mathcal{I} < X$.

L'avantage du concept d'accessibilité adiabatique est qu'il ne repose pas sur la définition de concepts tels que le chaud, le froid, la température, la chaleur, ni même sur une distinction entre chaleur et travail. La grande question est alors de savoir s'il est possible de définir la relation $<$ par une fonction ordinaire à valeurs réelles S , de sorte que si X et \mathcal{I} sont reliées par $<$, alors $S(X) \leq S(\mathcal{I})$ si et seulement si $X < \mathcal{I}$. En définitive, ce qui est visé par cette approche, qui est influencée par les travaux Carathéodory (1909) et Giles (1964), est un principe entropique à l'effet que l'entropie doit augmenter lors d'un processus irréversible (ce qui sera présenté plus loin). (Pour d'autres présentations axiomatiques du concept d'entropie, voir Serrin 1979 ; Zanchini & Beretta 2008.)

4.2.4.1 PROPRIÉTÉS

La fonction d'état $S_{\mathcal{I}}$ possède plusieurs propriétés qu'il est nécessaire de souligner afin de mieux cerner la signification du principe entropique et dans une certaine mesure du second principe de la thermodynamique. En définitive, c'est la stratégie explicative de la physique au moyen de ce concept qu'est l'entropie qui est en jeu. Il a été vu que ce concept était défini par une fonction d'état s'appliquant aux états d'équilibre et opérant une distinction entre processus réversible et irréversible.

1. L'entropie n'est *pas* une *propriété conservative*. Il n'y a donc pas une telle chose que la loi de la conservation de l'entropie. Plus précisément, l'entropie n'est conservée que dans le cas particulier des processus réversibles. Mais, tel que discuté, un processus réversible est considéré comme une idéalisation, un concept limite qui n'est jamais vraiment réalisée, à l'instar d'un mouvement rectiligne uniforme dans la mécanique newtonienne.

2. La fonction S est considérée comme *additive* et *extensive*, en ce sens que l'entropie d'un système composé de plusieurs parties correspond à la somme des entropies des parties⁴³. Cela est vrai si l'énergie du système composé correspond à la somme des énergies des parties et si la quantité de travail produite lors d'une transformation quelconque correspond aussi à la somme des quantités de travail

⁴³ Pour un système A , composé de deux sous-systèmes A_1 et A_2 tel que $A = A_1 + A_2$, l'entropie est additive si $S(A) = S(A_1, A_2) = S(A_1) + S(A_2)$.

produites par les parties. Si ces conditions sont remplies, ce qui est loin d'être toujours le cas, alors il est possible de déterminer l'entropie d'un système même si celui-ci n'est *pas* à l'équilibre. Et si ces conditions sont remplies, un système composé présentant la même entropie pour deux états différents peut être transformé d'un état à l'autre par un processus réversible sans d'autre effet dans l'environnement de ce système.

3. Comme il a été vu précédemment, puisque l'entropie thermodynamique est une variable définie par une fonction d'état, le changement d'entropie d'une substance quelconque passant d'un état à un autre est la même pour *tout* processus, qu'il soit réversible ou irréversible. Autrement dit, la différence d'entropie entre deux états d'équilibre est indépendante de la manière ou du processus ou du chemin pris pour passer d'un état à l'autre. En termes clairs, si l'entropie d'un état x_i est $S(x_i)$ et celle d'un état x_f est $S(x_f)$, alors la différence d'entropie $S(x_i) - S(x_f)$ est la même peu importe le processus reliant les états x_i et x_f . Mais il faut encore une fois préciser que cette différence d'entropie est égale à l'intégrale définie par $\int \delta Q/T$ si et seulement si le processus est réversible. En revanche, si le processus n'est pas réversible, la différence d'entropie entre l'état initial et l'état final n'est définie que par une borne inférieure, un seuil de valeur minimale.

4. Le calcul de l'intégrale $\int \delta Q/T$ demande que soit connue la relation entre Q et T pour un certain processus. Or, cette relation n'est pas toujours disponible et c'est pourquoi le calcul ne peut être effectué que dans certaines situations particulières. L'une de ces situations est celle d'un processus réversible où l'augmentation de la température est relativement faible, de sorte que la quantité δQ soit déterminée au moyen de la chaleur spécifique de la substance pour une plage de températures données⁴⁴. Il est alors possible de construire des diagrammes présentant la relation entre la température et l'entropie pour une substance donnée suivant un processus réversible.

5. La fonction d'état S ne permet pas de calculer les valeurs absolues de l'entropie, seulement les différences d'entropie entre deux états d'équilibre. Mais, puisque ces différences sont les mêmes pour tout processus, par définition même de ce qu'est une fonction d'état, il est possible de choisir un état d'équilibre arbitraire x_s par rapport auquel l'entropie d'un état d'équilibre quelconque est définie. Si le choix s'arrêtait sur un autre état d'équilibre arbitraire, alors la valeur de l'entropie ne différerait que d'une constante additive. C'est pourquoi l'entropie d'un *processus* entre un état x_f et cet état arbitraire x_s peut être considérée comme l'entropie de l'état x_f . De fait, tel que mentionné, selon le troisième principe de la thermodynamique, l'entropie de toute substance dite pure est considérée comme nulle au zéro absolu et peut donc être considérée comme un état de référence.

⁴⁴ $\delta Q = C dT$ et donc $\int [T_1 \rightarrow T_2] C dT/T = S(x_i) - S(x_f)$, où C est la chaleur spécifique de la substance et T sa température.

Les principales propriétés de la fonction S_T définissant l'entropie thermodynamique sont donc celles d'une fonction d'état définie pour des états d'équilibre reliés par un processus réversible d'après l'intégrale $\int \delta Q/T$. La condition d'équilibre assure que les variables de cette fonction, au premier chef la température, sont bien définies au sens où elles sont déterminées en tout point du système et déterminables expérimentalement. Cette fonction est par ailleurs bien définie lorsque la relation entre Q et T est établie, ce qui n'est pas toujours le cas. Mais si ces conditions sont remplies, la fonction représente alors un cas limite idéalisé d'un processus réversible reliant deux états d'équilibre pour lequel la différence d'entropie est établie. Dans le cas particulier d'un processus réversible et cyclique, cette différence est, par définition d'une fonction d'état, nulle. Dans tous les autres cas, pour les systèmes pouvant faire l'objet d'une éventuelle corroboration et suivant donc un processus qui n'est pas réversible, la différence d'entropie entre l'état initial et l'état final d'un tel processus est supérieure à la valeur de l'intégrale $\int \delta Q/T$.

4.2.5 Principe entropique

La fonction d'état S_T de Clausius a permis d'établir un nouveau concept physique ainsi qu'une nouvelle variable d'état (Chapitre 3). Elle ouvre ainsi la voie à une formulation mathématique du second principe de la thermodynamique. De sorte que les formulations précédentes, faisant appel à l'impossibilité de certains processus impliquant des quantités de chaleur et de travail, se voient remplacées par une formulation en fonction de cette nouvelle variable d'état. Si la rigueur qu'offre les mathématiques est souhaitable, il y a néanmoins un glissement ici qui s'opère dans l'appréhension empirique de la nouvelle formulation qui aboutit à plus d'ambiguïté quant à sa signification et sa portée.

La fonction d'état définissant l'entropie permet ainsi une formulation mathématique du second principe de la thermodynamique. Il a été vu que ce principe peut s'exprimer par la ST, qui stipule en termes simples qu'une quantité de chaleur ne peut être convertie complètement en travail. Cela revient à dire qu'il n'existe aucun processus présentant cette propriété, cette capacité. Suivant une logique sans doute héritée de la mécanique, où certaines équations déterminent l'évolution temporelle de l'état des systèmes, la traduction envisagée vient plutôt dire que l'entropie caractérisant l'état d'un système, dans certaines circonstances, ne peut qu'augmenter. Cela revient à dire qu'un système ne peut évoluer que vers des états compatibles avec une augmentation d'entropie. C'est ce qu'exprime en substance le *principe entropique*.

4.2.5.1 CARACTÉRISATION

Pour *tout* processus, $dS > 0$ s'applique aux processus irréversibles (et supposément « réels »), $dS = 0$ s'appliquent aux processus réversibles (idéalisés), et $dS < 0$ ne s'applique pas (du moins dans le cas d'un système isolé, ce qui est supposément le cas de l'univers comme un tout). La *génération* ou variation *totale* ou *globale* d'entropie, dS , peut être scindée en deux, soit en une *production* d'entropie ou entropie *interne*, $d_i S$, et un *flux* d'entropie ou entropie *externe*, $d_e S$. Tandis qu'une production d'entropie, qui est un accroissement d'entropie dû aux modifications intérieures du système, est strictement positive, un flux d'entropie, dû aux échanges entre le système et l'extérieur, c'est-à-dire son environnement, peut être négatif pour un système qui est alors ouvert, de sorte que $d_e S < 0$, ce qui correspond à une diminution *locale* d'entropie, c'est-à-dire pour le système en question. Mais une variation totale d'entropie, dS , ne peut être négative au sens où un changement total d'entropie incluant celui du système et celui de l'environnement ne peut pas être négatif. De sorte que $dS < 0$ est impossible, ce qui correspondrait à une diminution globale d'entropie. Aussi, $d_e S < 0$ est impossible, autrement une diminution locale d'entropie serait possible pour un système isolé. En d'autres mots, il est possible de réduire l'entropie d'un système par le truchement d'un autre système, mais l'entropie combinée de ces deux systèmes ne peut jamais décroître. En effet, le premier système est considéré comme étant ouvert, tandis que le système total combiné est considéré isolé. En somme :

$$dS = d_i S + d_e S \geq 0$$

(Prigogine 1968 ; Coveney 1988 ; Gnaiger 1993).

L'argument peut ainsi être généralisé à l'univers, considéré comme un système isolé et composé d'un système quelconque *et* de son environnement vu comme le reste de l'univers. C'est pourquoi l'entropie de l'univers, entendu comme la somme de tous les systèmes possibles, augmente⁴⁵. Dans le cas d'un système isolé, $d_e S = 0$ par définition, et dans le cas d'un système fermé ou ouvert, $d_e S$ demeure indéterminée et dépend des conditions du système et de son environnement, de leur interaction. Aussi, de par les propriétés de la fonction S établies précédemment, pour un processus réversible $d_i S = 0$ et pour un processus irréversible $d_i S > 0$, de sorte que pour tout processus $d_i S \geq 0$. En conséquence, dS est nulle si et seulement si le système est isolé et subit un processus réversible⁴⁶.

⁴⁵ On peut toutefois objecter que l'entropie de l'univers n'est pas un concept clair car on ne voit pas comment l'univers pourrait recevoir de la chaleur. Mais, ici, l'univers représente un système isolé. (Voir Clausius 1865 et Chapitre 3#.)

⁴⁶ On peut affirmer en fait que ces deux conditions ne sont *jamais* respectées puisqu'elles correspondent toutes deux à des idéalisations.

Ainsi, l'inégalité $\int dS = S(x_f) - S(x_i) \geq \int \delta Q/T$ rappelle que la production d'entropie est toujours supérieure au flux d'entropie ; ou encore, le changement d'entropie *total* est toujours supérieur ou égal à 0. De sorte que, si le flux d'entropie est nul, le changement d'entropie total est égal à la production d'entropie. En somme,

$dS > 0$ [processus irréversible],

$dS = 0$ [processus réversible],

$dS < 0$ [processus impossible].

L'expression $dS \geq 0$, qui équivaut à $S(x_f) \geq S(x_i)$, correspond au *principe entropique*. Il est souvent admis (par ex. Tuckerman 2010 : 60), du moins du point de vue macroscopique (Van Wylen, Sonntag & Desrochers 1992 : 221), que ce principe *est* le second principe de la thermodynamique. Mais cette identification ne va pas de soi et mérite de plus amples précisions.

Cette formulation du principe entropique est équivalente à celle de Lieb & Yngvason (1999 ; 2003 ; voir aussi Thess 2011 : 48) qui stipule que pour un processus adiabatique, tel que défini précédemment par la relation \prec , l'entropie ne peut pas décroître, c'est-à-dire que l'entropie de l'état final d'un processus est toujours supérieure ou égale à l'entropie de l'état initial. Et si l'entropie finale est supérieure à l'entropie initiale, alors le processus adiabatique est considéré comme irréversible. L'idée est qu'un système composé d'un système thermodynamique quelconque *et* de son environnement subit *toujours* d'une façon ou d'une autre un processus adiabatique, c'est-à-dire que l'échange de chaleur entre le système et son environnement ne fait pas varier la quantité de chaleur du système global (composé), donc l'entropie totale finale est toujours supérieure à l'entropie totale initiale $S(x_f) \geq S(x_i)$.

Le principe entropique présente ainsi plusieurs subtilités qui méritent qu'on s'y attarde. D'abord, il faut bien remarquer que, malgré l'affirmation à l'effet qu'il y a toujours augmentation de l'entropie *totale*, une diminution *locale* d'entropie demeure possible. Et cela est en effet observé dans les systèmes ouverts uniquement car, par définition, l'entropie d'un système isolé ne peut qu'augmenter ($d_i S \geq 0$). Si une diminution locale d'entropie d'un système isolé était possible, c'est-à-dire si une diminution *spontanée* d'entropie était possible, alors une diminution d'entropie *totale* serait possible, ce qui contredirait le principe entropique. Ensuite, la variation d'entropie (dS) est déterminable que si le processus est réversible et si la valeur de l'intégrale ($\int \delta Q/T$) est bien définie ; dans ce cas, $\int dS = S(x_f) - S(x_i) = \int \delta Q/T$. Ainsi, dans le cas particulier d'un processus réversible et adiabatique, la valeur de cette intégrale est nulle et la variation d'entropie aussi, de sorte $S(x_f) = S(x_i)$; on parle alors de processus ou d'évolution *isentropique*. Enfin, si le processus n'est pas réversible, alors la valeur de cette intégrale ne permet pas de déterminer la variation d'entropie, mais seulement sa valeur limite inférieure, une

« valeur plancher » en quelque sorte. Tel que mentionné, la variation d'entropie est alors $\int dS = S(x_f) - S(x_i) > \int \delta Q/T$. Toutefois, cette inégalité peut être transformée en égalité, telle que $\int dS = S(x_f) - S(x_i) = \int \delta Q/T + \xi$, où ξ représente une quantité à déterminer. Puisque $\delta Q < TdS$, il s'ensuit que $dS = \delta Q/T + \xi$, et que la quantité ξ a les unités de l'entropie, soit celles de la chaleur par celles de la température (J/°K). Cette quantité est parfois appelée « production irréversible d'entropie ». De plus, comme un processus réversible peut se caractériser par une production maximale de travail, contrairement à un processus irréversible, ce dernier peut se caractériser par une « perte de travail potentiel » (voir la discussion précédente ; van Wylen, Sonnatg & Desrochers 1992 : 217).

Cette idée de perte était déjà présente dans le second principe, comme on l'a vu, mais aussi dans la notion de « dégradation d'énergie » (Bruhnes 1909). Elle a d'abord été importante pour des raisons pratiques, plus proches de la réalité des ingénieurs, qui y voyaient une mesure de la qualité de l'énergie, qui semble prendre une forme de moins en moins « exploitable », de moins en moins « utilisable », au point de nécessiter d'autres ressources matérielles. Par exemple, une quantité de charbon peut être brûlée afin de chauffer de l'eau, produire de la vapeur et ainsi actionner un piston effectuant un travail, mais cette chaleur dégagée est inévitablement dissipée dans l'environnement (à l'extérieur de la machine) et devient inutilisable, ce qui nécessite la consommation d'une *autre* quantité de charbon afin de maintenir cette capacité d'effectuer un travail. On voit donc que l'énergie est « utilisable », « exploitable », « pratique » lorsqu'elle peut être convertie en travail et que ces attributs ne s'appliquent pas en propre à l'énergie sous forme de chaleur. Cette notion de « travail utile » apparaît dans un très grand nombre de descriptions de phénomènes, du déplacement de charges dans un processus industriel à la synthèse de composés chimiques complexes tels que l'acide désoxyribonucléique.

Afin d'illustrer encore plus clairement cette idée un exemple est de mise. Soit un radiateur électrique dans une pièce fermée, et supposons que la quantité d'énergie sous forme d'électricité pour le faire fonctionner se trouve dans les limites de ce système (dans une batterie par exemple) ; si le radiateur est mis en fonction alors l'énergie électrique est convertie en chaleur. L'énergie totale du système avant et après cette mise en fonction est strictement la même, mais la quantité d'énergie disponible pour effectuer un travail a diminué et l'entropie a augmenté. En considérant cette quantité d'énergie disponible comme celle pouvant faire fonctionner un moteur électrique, un ordinateur, etc., il est relativement intuitif de voir que le système à l'état final, soit une pièce plus chaude, ne possède pas cette capacité. Il semble aussi évident que l'état initial ne peut être recouvré de l'état final, la chaleur ne pouvant ici être convertie en électricité. Encore une fois l'entropie est associée à l'irréversibilité au sens d'irré récupérabilité. Mais elle est aussi associée ici à la notion de *qualité* d'énergie, de manière assez pragmatique et intuitive, en ce sens que l'état final de l'exemple représente un état où l'énergie est moins « utile », moins « exploitable », moins « pratique », mais aussi, en général, moins chère. C'est

justement cette tendance de l'énergie à prendre une forme moins « utile », moins « exploitable », moins « pratique », etc. qui explique le fait que l'énergie doit être payée malgré sa conservation⁴⁷.

On voit donc que lorsque la quantité d'énergie disponible pour effectuer un travail diminue l'entropie augmente, et vice-versa. Le concept de quantité d'énergie disponible pour effectuer un travail, dans son acception générique, c'est-à-dire selon les divers flux disponibles (masse, chaleur, travail), est le concept d'*exergie*, qui est de plus en plus utilisé dans la littérature. L'exergie est définie comme le maximum théorique d'énergie pouvant être convertie en travail (énergie mécanique ou électrique). De sorte que la qualité d'énergie peut être mesurée par la proportion d'exergie pour un système donné. L'exergie est ainsi inversement proportionnelle à la génération d'entropie. Il n'y a pas donc d'intérêt particulier de s'attarder plus longtemps sur ce concept (pour plus de détails, voir Wall 1977 ; Bejan 2002 ; Shukuya & Hammache 2002).

Mais la distinction entre le second principe de la thermodynamique et le principe entropique est importante pour plusieurs raisons. D'abord, il n'est pas si évident qu'il y a équivalence selon le sens donné précédemment. Si le second principe permet de décrire l'évolution des systèmes en termes empiriques, immédiats ou opérationnels, comme une « donnée directe de l'expérience », ce n'est pas tout à fait le cas avec le principe entropique. Il peut être relativement simple de trouver des équivalents phénoménologiques du concept de travail, comme la notion d'effort, et de celui de chaleur, comme la notion de chaud, mais il faudrait faire preuve de beaucoup d'imagination pour procéder de même avec l'entropie. On pourrait peut-être soutenir que cette différence d'appréciation trouve sa source dans la distinction entre les variables intensives, comme la pression, lesquelles peuvent justement être « ressenties », et les variables extensives, comme l'entropie et le volume, lesquelles ne le peuvent pas, semble-t-il – mais je ne poursuivrai pas ici cet argument. Ensuite, les quantités auxquelles réfèrent ces deux concepts que sont la chaleur et le travail, qui sont au cœur du second principe, peuvent être mesurés, avec, par exemple, un thermomètre et un ressort, respectivement, mais ce n'est pas le cas avec l'entropie, du moins pas de manière aussi directe. Un appareil visant à mesurer l'entropie – un « entropièmètre » – serait inévitablement un assemblage d'appareils de mesure⁴⁸. Enfin, si le principe entropique vise une synthèse des différentes formulations du second principe de la thermodynamique, il

⁴⁷ Le lien avec les problèmes environnementaux et le développement durable, comme quoi le maintien de la qualité d'énergie dans une société industrialisée exige beaucoup de ressources, a déjà été évoqué. L'idée est qu'une « consommation » d'énergie, plutôt une transformation d'énergie, implique une dégradation d'énergie au sens où une quantité moindre de travail utile peut être extraite du système en question. Et cette capacité à effectuer un travail utile ne peut être localement recouvrée qu'au prix d'une autre consommation de ressources et d'énergie, bref en ouvrant les limites du système (qui n'a plus cette capacité d'effectuer un travail).

⁴⁸ Atkins (1984 : 40) donne un exemple d'un tel appareil, lequel consiste essentiellement en « une sonde comprenant une résistance chauffante et un thermomètre dont les indications sont traitées par un microprocesseur qui affiche la mesure de l'entropie ».

le fait en introduisant un nouveau concept – l'entropie – et en s'appuyant sur d'autres hypothèses. L'équivalence recherchée ne va donc pas de soi, ni l'affirmation que ces deux principes seraient en réalité synonymiques.

Alors, est-ce que le second principe de la thermodynamique et le principe entropique sont équivalents ? Il a été vu que le second principe recevait plusieurs formulations et c'est pourquoi la séquence thermodynamique (ST) est ici employée puisqu'elle peut résumer assez bien l'« essence » de ce principe⁴⁹.

Prenons un exemple simple : deux corps, l'un chaud et l'autre froid, mis en contact. Si l'on considère le système formé de ces deux corps, alors ce système peut être considéré comme isolé. Dans une telle situation, une quantité de chaleur Q va passer du corps chaud à une température T_+ au corps froid à une température T_- , de sorte que le corps chaud va céder une quantité d'entropie égale à $-Q/T_+$ et le corps froid va recevoir une quantité d'entropie égale à $+Q/T_-$. Considérant chaque corps pris isolément, ceux-ci constituent des systèmes ouverts puisque l'un cède de la chaleur tandis que l'autre en reçoit. C'est pourquoi le corps chaud voit son entropie diminuer, ce qui ne contredit pas le principe entropique puisqu'il s'agit d'un système ouvert, tandis que le corps froid voit son entropie augmenter. Mais l'entropie totale du système composé de ces deux corps, autrement dit, toute chose étant égale par ailleurs, l'entropie de l'« univers », augmente puisque la quantité $+Q/T_-$ est supérieure à la quantité $-Q/T_+$. Or si le transfert de chaleur avait été effectué dans le sens inverse, du chaud vers le froid, l'entropie totale du système composé de ces deux corps aurait diminuée puisque la quantité $-Q/T_-$ est supérieure à la quantité $+Q/T_+$, donc $Q/T_+ + -Q/T_- < 0$, ce qui est en contradiction avec le principe entropique. C'est aussi en contradiction avec l'énoncé de Clausius à l'effet que la chaleur ne peut passer du froid au chaud d'elle-même. C'est en effet une conséquence de l'entropie, qui est définie en fonction de la température et qui ne peut qu'augmenter, d'interdire ce type de transfert de chaleur.

Prenons un autre exemple : un corps chaud cédant de la chaleur à une machine produisant une quantité de travail. Si cette quantité de travail devait être égale ou supérieure à la quantité de chaleur reçue, alors, il s'agirait clairement d'un cas de contradiction avec l'énoncé de Kelvin ainsi qu'avec la ST. Dans une telle situation, une quantité de chaleur Q_+ est cédée du corps chaud à une température T_+ de sorte que le corps chaud va céder une quantité d'entropie égale à $-Q_+/T_+$. Mais comme la machine effectue un travail sur l'environnement, il n'absorbe pas de chaleur en raison du premier

⁴⁹ Cette affirmation peut (encore) porter à polémique. On ne peut que se référer à l'argumentation précédente. Mais je souhaite y ajouter un élément dans le contexte de la présente discussion : la ST peut ici se « limiter » au principe de l'impossibilité du mouvement perpétuel de second type, de sorte que plusieurs éléments que la ST « ajoute » ne sont pas nécessaires dans la discussion qui suit.

principe et de l'hypothèse de départ, et son entropie ainsi que celle de l'environnement demeurent inchangées. En conséquence, l'entropie totale du système composé du corps chaud et de la machine aurait diminuée d'une quantité égale à $-Q_+/T_+$. Selon le principe entropique cette diminution doit être compensée par une quantité d'entropie supérieure : il doit y avoir « surcompensation ». Cela est fait par une quantité d'entropie égale à $+Q_-/T_-$ et cédée à l'environnement ou au réservoir froid, ce qui équivaut à cette « taxe » payée pour la conversion de la chaleur en travail.

Le principe entropique est donc violé dans les mêmes conditions que les énoncés du second principe de Clausius et de Kelvin. Toutefois, cette affirmation n'est pas tout à fait juste. En effet, ces conditions demandent que les différents concepts impliqués dans ces deux principes soient bien définis. Car il se peut que dans certaines conditions l'équivalence soit clairement établie, mais que dans d'autres conditions, seulement l'un de ces principes soit bien défini et que l'autre ne le soit pas. Ainsi, les ensembles respectifs circonscrivant les situations où ces principes sont bien définis et pertinents du point de vue de la thermodynamique ne seraient pas coextensifs. Par conséquent, il y aurait des situations où l'un de ces principes serait bien défini et pertinent, mais où l'équivalence ne serait pas démontrée. Par exemple, l'entropie d'un système complexe peut être indéfinie tandis que les concepts de chaleur et de travail puissent tout de même s'y appliquer et qu'au surplus il soit impossible d'y harnacher un mouvement perpétuel. Ainsi, on peut bien imaginer des situations où le principe entropique ne s'applique pas mais où le principe de l'impossibilité du mouvement perpétuel et/ou le second principe seraient valides. C'est pourquoi, d'une part, la validité des critiques à l'endroit du principe entropique n'implique pas la possibilité d'obtenir un mouvement perpétuel de second type. C'est aussi pourquoi, d'autre part, la démonstration de l'équivalence entre le second principe et le principe entropique n'est pas aussi simple qu'on voudrait bien le croire. Si les arguments précédents sont souvent considérés comme suffisants, l'équivalence devrait en toute rigueur être ramenée à quelque chose d'aussi fort qu'une réduction théorique (en supposant que les conditions d'une telle réduction soient bien définies).

Les deux derniers exemples peuvent aussi servir à établir l'équivalence du principe entropique avec la ST. Un tel exercice peut sembler superflu si l'on considère l'argumentation précédente quant à l'équivalence de celle-ci avec les énoncés du second principe de Clausius et de Kelvin comme valide. Mais voici brièvement de quoi il retourne. La ST dit qu'une différence de température entraîne un flux de chaleur du chaud vers le froid, étant donné qu'un flux de chaleur du froid vers le chaud exige une consommation de travail, et que ce flux de chaleur peut produire une quantité de travail moindre que la quantité de chaleur transférée. Puisque le premier principe n'est pas remis en cause par la ST, il faut donc que l'environnement reçoive une partie de cette chaleur, autrement il n'y aurait pas équivalence entre travail et chaleur ou une quantité de chaleur serait annihilée. Puisque cette quantité de chaleur est transmise à une température moindre, on retrouve donc l'augmentation d'entropie du principe

entropique tel qu'illustrée dans l'exemple précédent. Inversement, la ST dit qu'un travail peut entraîner un flux de chaleur et créer ou intensifier une différence de température. Dans ce cas, en mode pompe à chaleur ou réfrigérateur, le corps froid cède des quantités de chaleur et d'entropie inférieures à celles reçues par le corps chaud. Car en vertu du premier principe, la quantité de travail nécessaire pour ce transfert est égale à la différence entre la chaleur reçue et celle cédée par le système, et la chaleur cédée doit ici être supérieure à celle reçue, autrement un mouvement perpétuel serait alors en puissance. On peut donc considérer, dans ces conditions simples, la ST comme équivalente au principe entropique.

4.2.5.2 CRITIQUES

Bien sûr, plusieurs objections ont été adressées au principe entropique. Bien souvent les objections concernent les manques d'une argumentation visant à asseoir la prétention d'une portée très générale, voire universelle, de ce principe. Cette prétention est sans doute motivée à la base par la robustesse du principe de l'impossibilité du mouvement perpétuel. Tout ce qui serait vraiment prouvé, objecte-on, ne s'appliquerait qu'aux états d'équilibre, qu'aux gaz parfaits, qu'aux fluides, qu'aux processus sans réaction chimique, etc. La portée de ce principe serait ainsi très limitée et ne permettrait donc pas d'expliquer les phénomènes qu'il visait au départ. Un tel résultat serait donc opposé à la motivation initiale du principe de l'impossibilité du mouvement perpétuel, le cas échéant. La question se pose donc de savoir ce qui mène à cet écart entre l'objectif et le résultat. La réponse se situe, je crois, dans le décalage qui existe souvent entre les diverses argumentations et leurs objections respectives quant à ce qui est raisonnable de considérer comme des quantités négligeables ou des *idéalisations*⁵⁰. Par exemple, certaines d'entre elles considèrent les effets sur l'environnement d'une connexion diathermique comme négligeables ou encore qu'un travail peut s'effectuer sans aucune friction. Sans les mêmes hypothèses, nul doute que les conclusions divergent. Mais le point d'achoppement le plus crucial est sans doute le suivant : l'hypothèse d'un processus cyclique réversible est nécessaire à la définition d'une fonction d'état, mais elle ne peut être qu'une idéalisation. Voyons ce que cela implique.

Reprenons d'abord brièvement l'argumentation de l'entropie thermodynamique. L'hypothèse d'un processus cyclique réversible permet d'établir ce que Clausius appelle des « valeurs d'équivalence », telle que $\sum f(T_i)Q_i = 0$, et si ce processus suit un cycle de Carnot alors on peut montrer que $f(T_i) = 1/T_i$, de sorte $\sum Q_i/T_i = 0$. Ce raisonnement permet alors de définir une fonction d'état – l'entropie : $\sum Q_i/T_i \approx \int \delta Q/T = \int dS$. Mais comme un processus cyclique réversible est un cas limite idéalisé, il s'ensuit

⁵⁰ On retrouvera ce genre de problème dans le prochain chapitre avec le « paradoxe du démon de Maxwell ». (Voir aussi Chapitre 2.)

qu'aucun processus ne respecte strictement cette condition. Donc, trivialement, tout processus « se produisant dans la nature » ne respecte pas cette condition. Ainsi, pour tout processus passant d'un état x_i à un état x_f , l'entropie est définie telle que $S(x_f) > S(x_i) + \int \delta Q/T$, autrement dit la valeur de l'entropie de l'état final ne peut être complètement déterminée à partir des valeurs de l'entropie de l'état initial et de l'intégrale. La fonction d'état ne définit donc qu'un seuil de valeur, un « plancher », de la variation d'entropie entre deux états. Il paraît alors naturel que certaines objections affirment que la portée de l'entropie thermodynamique est dérisoire puisqu'elle ne s'appliquerait à aucun processus « réel », ou encore que sa démonstration ferait appel en réalité à des processus qui ne sont pas réversibles, en conséquence de quoi la fonction d'état ne pourrait être établie.

Encore une fois, une comparaison avec la première loi de la mécanique classique, soit la loi de l'inertie, est éclairante. Cette loi stipule qu'un corps libre, sans force s'y appliquant, persiste dans un état de mouvement rectiligne uniforme. Or un mouvement rectiligne uniforme est un cas idéalisé jamais instancié. On peut ainsi inférer qu'un corps qui ne suit pas un mouvement rectiligne uniforme suit un mouvement influencé par une force quelconque, laquelle est à l'origine, en est la « cause » (un concept polysémique), et donc explique pourquoi le corps n'est pas dans un état de mouvement rectiligne uniforme. En comparaison, la thermodynamique identifie aussi des « causes » visant à expliquer pourquoi le système ne suit pas un processus réversible. Trivialement, ces « causes » sont des irréversibilités. Généralement, il s'agit d'un échange de chaleur qui n'est pas accompagné d'une production de travail. Par exemple, une irréversibilité peut correspondre à la chaleur dégagée par friction. Typiquement, lors d'un tel processus l'entropie augmente. Le défi consiste alors à traduire les irréversibilités en termes d'entropie afin de l'articuler au sein d'une explication pertinente fournie par le principe entropique. C'est-à-dire, pour que le bilan entropique soit positif et ainsi respecter ce principe, il faut tenir compte de l'ensemble des échanges et productions d'entropies, ce qui n'est pas toujours évident (par ex. les traces de pneus laissées par une voiture serait de l'entropie). Bien que le principe entropique puisse être appliqué dans bon nombre de situations, il semble clair que la thermodynamique orthodoxe ne fournit pas les outils nécessaires pour une telle traduction devant la variété des situations où il y a variation d'énergie par échange de chaleur et de travail (section 4.2.7).

Par ailleurs, la stratégie visant à définir une fonction d'état a bien entendu un impact sur la stratégie explicative du principe entropique (soit : l'entropie totale doit augmenter malgré des diminutions locales). Et une comparaison avec la première loi de la mécanique classique est ici aussi utile car dans les deux cas le changement d'état semble être expliqué par rapport à un cas limite idéalisé. Les discussions à propos des idéalizations (Chapitre 2) dans les modèles et lois scientifiques sont nombreuses et s'accordent en général pour dire qu'elles n'ont pas constitué un obstacle majeur au progrès

scientifique, ou dit autrement, les idéalizations ne sont intrinsèquement illégitimes en science (ce point sera discuté plus loin).

Une autre critique importante du concept d'entropie concerne la notion d'*équilibre*. En général, et tel qu'indiqué plus tôt, un système thermodynamique est dit être dans un état d'équilibre si ses propriétés sont stables dans le temps et peuvent être décrites par un petit nombre de quantités macroscopiques. Un état d'équilibre est aussi considéré comme un état indépendant du temps au sens où il demeure inchangé mais aussi au sens où il ne dépend pas de l'histoire passée du système. Il s'agit d'une notion essentielle en thermodynamique. Car la condition que les variables ou paramètres soient macroscopiques et qu'elles réfèrent à des états d'équilibre constitue le principal « critère axiomatique » distinguant les variables thermodynamiques de celles qui ne le sont pas (Landsberg 1956 : 365). Il s'agit en d'autres termes d'une condition essentielle pour que la thermodynamique soit autonome. En outre, l'existence d'un état d'équilibre, descriptible par quelques paramètres macroscopiques dont les valeurs et les interrelations ne changent pas dans le temps, est le « fait thermodynamique le plus fondamental » auquel la mécanique statistique doit rendre compte (Sklar 1993 : 156). C'est pourquoi l'équilibre est généralement *défini* comme une situation où aucune transformation n'est possible sans intervention extérieure, où les paramètres décrivant l'état du système sont constants dans le temps. Il y a plusieurs types d'équilibre, l'équilibre mécanique où la pression est constante, l'équilibre électrique où il n'y a pas de différence de potentiel, l'équilibre des espèces où la concentration des espèces demeurent inchangée, etc. Dans le cas de l'équilibre thermique, par exemple, celui-ci est défini comme une situation où deux corps mis en contact ne donnent lieu à aucun changement observable préalablement déterminé, comme l'expansion d'un volume de mercure par exemple. La conclusion découlant d'une telle situation est que ces deux corps partagent une même propriété qui est celle d'avoir la même température.

Ce « critère axiomatique » auquel Landsberg fait référence s'appuie sur deux constats empiriques. Ces derniers déterminent non seulement ce « fait physique », érigé par induction en « critère axiomatique », mais aussi des critères opératoires permettant d'identifier dans les faits une situation particulière d'équilibre. De manière analogue, on peut tirer par induction l'affirmation générale « les corps tombent », mais cette affirmation peut aussi être l'objet d'une définition opératoire de la gravité permettant l'identification dans les faits d'un cas particulier de manifestation de la gravité. Le premier constat est celui de la constance, « après un certain temps », dans les paramètres caractérisant un système, ceux-ci étant alors indépendants du temps. Il s'agit bien là de la définition opératoire de l'équilibre. Mais cette inférence est similaire à celle passant d'une « absence de preuves » à une « preuve d'absence » : l'observation de certains paramètres constants mène à la conclusion qu'il y a nécessairement absence de changement dans tout le système. Une telle inférence inductive visant à identifier une situation d'équilibre dans les faits est bien sûr invalide. C'est pourquoi, un second constat

empirique est nécessaire, celui à l'effet qu'il n'y ait pas d'interaction entre le système et son environnement. D'ailleurs, face à la question de l'équilibre, ces interactions sont généralement appelées des « perturbations », entendues comme ce qui cause un changement dans le système. La combinaison de ces deux observations ou constatations appuierait l'inférence à une situation d'équilibre, laquelle va comme suit : d'abord, l'équilibre est défini de manière opératoire comme l'absence de changement dans certains paramètres définissant l'état du système ; ensuite, l'observation n'apporte pas de preuve ni de changement dans ces paramètres ni de cause possible de changement ; donc, le système est à l'équilibre. Cette inférence est aussi portée plus loin puisque la simple constatation de l'absence d'interaction est parfois jugée suffisante pour inférer la situation d'équilibre. De fait, l'expérience atteste qu'après « un certain temps » les systèmes atteignent un état d'équilibre en l'absence d'interaction.

Mais en fait, et c'est là l'objection, très peu de systèmes, pour ne pas dire aucun, se trouvent dans un état d'équilibre strict, où les deux critères susmentionnés, la constance des paramètres décrivant l'état du système et l'absence d'interaction avec l'environnement, sont respectés. Ainsi il est rare qu'il n'y ait aucune différence de potentiel (ou potentiel déséquilibré) ni forces motrices en jeu ; on peut penser à la gravité, à des différentiels de pression ou de potentiels chimiques ou thermiques. Un système parfaitement isolé de son environnement semble ainsi être une idéalisation, un cas limite jamais instancié. On ne peut pas, par exemple, aménager un bouclier de protection contre la gravité. De plus, le temps que met le système à atteindre cet état d'équilibre, appelé « temps de relaxation », peut être de quelques microsecondes à plusieurs millénaires. Il n'est donc pas évident de savoir si un système particulier a effectivement atteint un état d'équilibre ou non. Ces difficultés témoignent de la notion d'équilibre (mais aussi d'homogénéité) comme relevant en grande partie d'une exigence pratique (Fermi 1936 : 1 ; Bridgman 1941 : 75). De sorte que la condition d'équilibre est souvent tenue pour acquise, autrement la thermodynamique n'aurait pas les outils nécessaires pour rendre compte des systèmes qui sont justement « thermodynamiques ». C'est pourquoi une telle définition peut sembler « circulaire » (Callen 1985), car un système est dans un état d'équilibre si ses propriétés peuvent être décrites par la thermodynamique.

Cependant, le concept d'équilibre peut représenter avec une bonne approximation bon nombre d'états observés. D'une part, les changements observés peuvent être considérés comme « négligeables », c'est-à-dire que les variations dans les paramètres caractérisant un système considéré à l'équilibre peuvent être de beaucoup inférieures à la valeur de ces mêmes paramètres. D'autre part, la validité de cette approximation, dans certains contextes, découle du fait que la notion d'état d'équilibre peut se confondre avec celle d'état tout court. Pourquoi ? Parce que le raisonnement décrit plus tôt fait état des éléments opératoires de la définition d'équilibre. En effet, un état qui n'est pas à l'équilibre peut être difficile à déterminer, en pratique comme en théorie, c'est-à-dire qu'il peut être difficile de mesurer

certaines paramètres le caractérisant ou d'inférer certaines de ces caractéristiques, qui varient dans le temps et dans l'espace, à partir des paramètres effectivement mesurés. D'où cette exigence pratique. Or on ne saurait prétendre qu'un système hors d'équilibre ne possède pas d'état. Que celui-ci soit changeant, on le conçoit, mais il est censé tout de même exister, et ce, malgré nos limitations perceptuelles, expérimentales, métrologiques ou calculatoires. Enfin, un système qui n'est pas à l'équilibre est typiquement en interaction avec son environnement. Étant donné les constantes variations de ses caractéristiques, il devient très ardu d'établir ces mêmes caractéristiques et par le fait même les limites du système lui-même. Car ces limites ne sont pas nécessairement fixes dans l'espace et il est donc difficile de départager le système de son environnement. Ainsi, un système qui n'est pas à l'équilibre interagit fortement avec son environnement et les variables qui caractérisent son état sont fluctuantes. Ce n'est donc pas que le système n'a pas d'état, ce qui serait absurde, mais plutôt que cet état est difficilement déterminable, donc largement et effectivement indéterminé⁵¹.

Un autre problème similaire découle de l'indétermination, dans une situation de non-équilibre, des paramètres entrant dans la définition de l'entropie, au premier chef la température, de sorte que le calcul de l'intégrale $\int \delta Q/T$ et la définition même de cette fonction d'état sont en péril. Car dans une telle situation la température n'est pas la même en tout point du système. Mais ce problème touche davantage la détermination expérimentale ou calculatoire de la température d'un système qui n'est pas à l'équilibre que sa définition. Si la température est définie comme l'énergie cinétique moyenne des molécules, alors il semble tout à fait raisonnable de dire qu'un système qui n'est pas à l'équilibre possède une température, la question étant plutôt de déterminer qu'elle est sa valeur. Et si cette condition de l'équilibre thermique n'est pas satisfaisante, si la température d'un système isolé varie d'un point à l'autre, on ne peut pas en conclure que le principe entropique est violé. Mais une corroboration expérimentale ou une falsification demeurent difficiles. De manière similaire, ce n'est pas parce la détermination de la valeur d'énergie exacte d'un système présente des difficultés en pratique qu'on en conclut que le principe de conservation de l'énergie est violé. Aussi, la condition d'équilibre est souvent assimilée à celle d'un processus quasi-statique. La raison en est relativement simple. Dans la construction de la fonction d'état S , les systèmes impliqués échangent de la chaleur avec des réservoirs thermiques, ce qui implique bien sûr une différence de température. Comme cette construction demande un processus réversible afin d'obtenir les « valeurs d'équivalence », et puisque la chaleur ne peut passer « d'elle-même » du froid vers le chaud, l'évolution réversible d'échange de chaleur est

⁵¹ Cette situation est similaire à ce que l'on retrouve en mécanique quantique où un système est indéterminé puisqu'il est considéré comme étant dans une « superposition d'états ». Mais, dans ce cas, il est commun, quoique discutable, d'y ajouter un élément ontologique, en ce que le système serait « fondamentalement » indéterminé sans égards à nos capacités de mesure, de calcul, etc., permettant de connaître cet état.

définie comme celle où l'échange de chaleur procède d'une différence de température infinitésimale. Il s'agit bien sûr, ici aussi, d'une idéalisation, car l'échange d'une quantité non-négligeable de chaleur par le truchement d'une différence de température infinitésimale demanderait un temps infini. La condition d'équilibre de même que la condition d'un processus quasi-statique relèvent donc d'impératifs principalement pratiques puisqu'il s'agit aussi d'idéalisations.

C'est pourquoi une éventuelle objection à l'effet que le principe entropique ne renvoie à rien de « réel » ou encore que ce qu'il cherche à expliquer est absent de la définition du concept d'entropie, risque de s'attaquer *in fine* à une stratégie explicative faisant intervenir des idéalisations (Chapitre 2). Par ailleurs, si l'équilibre est une condition nécessaire à la définition de l'entropie, on voit mal comment celle-ci pourrait fournir une explication de la tendance à l'équilibre des systèmes thermodynamiques. Comment, en d'autres mots, une hypothèse fondamentale du principe entropique peut-elle en être une conséquence ? Cette possibilité est évidemment à rejeter. Malgré ces critiques, le principe entropique peut recevoir des appuis indirects venant de l'expérience. Ainsi, si ce principe est une conséquence de l'impossibilité du mouvement perpétuel de second type, alors la non-falsification de cette dernière est aussi un appui au principe entropique. De plus, bien que ni le succès d'un principe ni son utilité n'établissent sa vérité, le principe entropique permet des prédictions fiables. Mais son appui le plus important, du moins aussi important que celui de l'impossibilité du mouvement perpétuel si l'on admet la validité de leur lien logique, provient des théories visant à « surpasser » la thermodynamique classique en prenant en compte justement des phénomènes irréversibles et des situations de non-équilibre. En effet, celles-ci prennent généralement pour acquise la validité du concept d'entropie (voir la prochaine section).

En conclusion, rappelons que le principe entropique stipule que l'entropie ne peut pas décroître. Cela signifie que pour tout système thermodynamique, défini à l'équilibre et plus spécifiquement avec une température déterminée, tout processus ne peut mener qu'à une augmentation de l'entropie totale, incluant celle du système en question et de son environnement. Plus précisément, pour tout processus passant d'un état x_i à un état x_f , le principe entropique veut que $S(x_f) \geq S(x_i) + \int \delta Q/T$, c'est-à-dire que la valeur de l'entropie de l'état final est supérieure ou égale à la valeur de l'entropie de l'état initial plus la valeur de l'intégrale définie par la quantité de chaleur échangée par le système divisée par la température à laquelle s'effectue cet échange. Lorsque le processus est réversible, l'équation précédente devient une égalité telle que $S(x_f) = S(x_i) + \int \delta Q/T$. Il s'agit cependant d'un cas idéalisé jamais instancié. Si le processus est cyclique, ou si le processus est adiabatique ($\delta Q = 0$), alors l'intégrale $\int \delta Q/T$ devient nulle et on a donc $S(x_f) \geq S(x_i)$.

Le principe entropique ainsi défini n'offre *prima facie* aucune assise à une interprétation comme « mesure du désordre », de notre « ignorance » ou comme une « probabilité » (ce qui sera discuté dans le prochain chapitre). En revanche, l'entropie peut être considérée comme une mesure des « irréversibilités » : plus la génération ou variation totale d'entropie est grande, plus grandes sont les irréversibilités. Cela signifie que plus la variation totale d'entropie, incluant celles du système et de son environnement, est grande, plus le processus auquel s'applique cette variation s'éloigne du cas limite idéalisé représentant un processus réversible où l'état initial du système peut être recouvré dans les limites de ce même système. C'est pourquoi un processus réversible définit la valeur limite inférieure de la variation totale d'entropie. Mais une caractérisation précise des « irréversibilités » devra sans doute attendre les développements d'une théorie prenant la pleine mesure des situations de non-équilibre et des processus irréversibles ; ce vers quoi nous nous tournons à l'instant.

4.2.6 Non-équilibre

Il a été vu que l'entropie thermodynamique d'un état x , $S(x)$, se définit comme l'intégrale de la quantité $\delta Q/T$ d'un état x_S vers cet état x selon un processus réversible, où x_S est un état arbitraire fixe. Puisque le processus est réversible, il s'ensuit que les états de ce processus sont des états d'équilibre. En conséquence, le processus est dit quasi-statique et les états initial et final doivent aussi être des états d'équilibre. Donc, l'entropie thermodynamique ne s'applique tout simplement pas aux états de non-équilibre. Ainsi, la thermodynamique – ou plutôt la thermostatique – n'a que très peu à dire sur un système avant qu'il n'atteigne un état d'équilibre. Cela ne veut cependant pas dire que la physique est impuissante à décrire ou expliquer les systèmes de non-équilibre. Mais, pour être en mesure de le faire, la physique doit se munir de l'équivalent du second principe pour les états de non-équilibre. La thermodynamique des processus irréversibles a tenté de le faire mais elle n'est valide que dans le voisinage immédiat de l'équilibre et ne prétend pas à plus de généralité. C'est pourquoi une théorie traitant des processus loin de l'équilibre est nécessaire. Les tentatives de solution ont généralement porté sur l'hypothèse que l'entropie peut être exprimée comme une fonction du temps. Mais cette hypothèse n'a reçu que très peu de justification. Il est pour le moins étrange qu'un problème d'une telle importance soit resté pour l'essentiel ouvert jusqu'à nos jours. Les discussions épistémologiques sur le non-équilibre, en thermodynamique comme en général, sont par conséquent plutôt limitées⁵².

⁵² Le livre édité par Hooker (2011), *Philosophy of Complex Systems*, est très varié, touchant le sujet de la complexité par exemple en anthropologie et en médecine chinoise, ce qui est en fait sa force et sa faiblesse, mais aussi il se veut davantage exploratoire, se bornant à présenter ce que *pourrait* être la philosophie de la complexité et par le fait même du non-équilibre ; dans le même

L'objectif ultime de la *thermodynamique des processus irréversibles* (TPI), comme la thermodynamique générique, est la détermination de différents champs représentant la densité de masse, la vitesse et la température en tout point du système. Cette détermination doit s'appuyer sur des relations reliant ces champs et les lois de conservation de la masse, des moments et de l'énergie interne. Il est supposé que le système peut être partitionné en différentes régions où les fluctuations microscopiques sont négligeables et que les équations fondamentales de la théorie dépendent des mêmes ensembles de variables qu'à l'équilibre – cette hypothèse fondamentale est appelée *principe d'équilibre local*. L'heuristique consiste alors à postuler une inégalité entropique (discutée plus loin) et de l'articuler dans l'équation de Gibbs de l'équilibre thermodynamique (thermostatique). Dans ce contexte, la clé de la manipulation du concept d'entropie consiste en l'artifice de sa représentation comme une substance capable de s'écouler dans l'espace telle un fluide. Cependant, l'entropie n'est *pas* une propriété qui se conserve puisqu'elle augmente pour tout système isolé. L'artifice consiste alors à modéliser l'entropie comme un fluide pouvant être détruit ou produit. Ainsi, de manière tout à fait semblable à ce qui a été présenté plus tôt, le taux de production d'entropie, soit la quantité d'entropie créée par unité de temps, correspond à la somme du taux d'accroissement d'entropie pour un volume donné et du flux externe d'entropie pour ce même volume.

La stratégie consiste alors à caractériser le système par un ensemble de flux et de forces, à l'origine de cet état de non-équilibre. Plusieurs équations de flux, à l'instar de l'équation de Fourier exprimant la conduction thermique, prennent la forme suivante : $\mathcal{J} = LX$, où X représente un gradient ou une affinité, aussi appelée « force généralisée », \mathcal{J} un flux ou un taux de réaction, et L est un coefficient indépendant, ce qui forme ce qu'on appelle les *équations phénoménologiques*. L'idée, similaire à celle articulant la force et l'accélération dans la mécanique newtonienne, est qu'un gradient ou une affinité induit un flux ou un taux de réaction, autrement dit que les forces généralisées « causent » les flux. La production d'entropie est alors déterminée par une somme de flux, positive ou nulle, et la situation où cette production est minimale correspond à un état stationnaire d'un ordre équivalent au nombre de paramètres maintenus fixes. Ainsi, à l'état stationnaire, l'activité du système augmente continuellement l'entropie du milieu, mais de la valeur minimale compatible avec les conditions aux limites définies par différents paramètres. Dans ce contexte, l'état d'équilibre n'est rien d'autre que l'état stationnaire particulier accessible lorsque les conditions aux limites admettent une production d'entropie nulle. Cela correspond ainsi au *principe de production minimale d'entropie* (voir Yourgrau, van der Merwe & Raw 1966 ; Prigogine 1968 ; Prigogine & Stengers 1979 ; Kondepudi 2008 ; Hutter & Wang 2010).

ton, mais sans doute moins systématique, voir aussi Waldrop (1992). Pour des discussions philosophiques sur le non-équilibre en thermodynamique, voir aussi Batterman (1991), Edens (2001), Bishop (2004 et 2011).

Le principal reproche que l'on peut adresser à cette approche concerne le principe d'équilibre local. En effet, on peut rester perplexe devant la prétention d'expliquer les phénomènes de non-équilibre avec un tel principe. On peut, dans un premier temps, remarquer que la température est définie de deux manières différentes, l'une comme une propriété continue à la frontière de deux corps et l'autre comme l'énergie cinétique moyenne des particules constituant un corps, par définition macroscopique. Dès lors on peut montrer que ces deux concepts de température diffèrent, excepté à l'état d'équilibre, et qu'en conséquence l'entropie de non-équilibre ne peut être équivalente à l'entropie d'équilibre. Dans un second temps, on peut montrer que le flux d'entropie n'est pas parallèle au flux de chaleur, mais qu'il dépend aussi du tenseur des contraintes, influencé par la viscosité de la substance. Cela ne remet pas en cause le principe entropique comme tel, mais impose tout de même des modifications dans le calcul du flux d'entropie. La conséquence est que ce calcul est quadratique plutôt que linéaire, comme c'est le cas en thermodynamique des processus irréversibles, ce qui permet d'éviter le problème de la propagation infinie de la chaleur et des contraintes en cisaillement.

Il a été vu (Chapitre 3) que ce paradoxe de la conduction thermique avait motivé le développement de la *thermodynamique prolongée* (« *extended thermodynamics* ») – une théorie développée principalement par Ingo Müller (voir Greven, Keller & Warnecke 2010). Celle-ci se distingue de la thermodynamique des processus irréversibles sur deux points. D'abord, elle généralise l'équation de Gibbs⁵³ de la thermostatique en posant l'hypothèse que l'entropie de non-équilibre dépend de quantités qui disparaissent à l'équilibre, comme le flux de chaleur et le flux de la quantité de mouvement. Vient ensuite l'hypothèse d'un flux d'entropie quelque peu modifié en ceci qu'il est aussi fonction de certaines quantités *constitutives*, soit des quantités spécifiques à une substance donnée approxinant la réponse de cette substance à certains stimuli extérieurs. À l'instar de la thermodynamique des processus irréversibles, l'objectif de la thermodynamique prolongée est la détermination d'un certain nombre de champs incorporés dans un vecteur de champs. La détermination de ces champs nécessite des équations de champ empruntées à la mécanique et à la thermodynamique. Ces équations déterminent à leur tour des relations constitutives respectant le critère de localité spatio-temporelle en ce que les vecteurs dépendent uniquement des valeurs des champs pour un événement donné. Chaque solution des équations de champ constitue un processus thermodynamique. Mais il n'y a aucun matériau pour lequel les relations constitutives sont connues. Le défi consiste donc à déterminer ces relations au moyen de trois outils disponibles, soit l'inégalité entropique, le critère de convexité et le principe de relativité.

⁵³ $dS = 1/T(dU + PdV)$.

L'hypothèse de départ pour la généralisation du concept d'entropie est de considérer les différents flux des équations constitutives comme des variables indépendantes de manière à les incorporer dans un ensemble de variables duquel dépendrait l'entropie. Les différentes propriétés de l'entropie généralisée sont les suivantes (Jou, Casas-Vàquez & Lebon 1988 ; Müller 2010). D'abord, comme précédemment, l'entropie est additive. Ensuite, l'entropie est définie par une fonction convexe des variables classiques mais aussi des flux, et la conséquence d'une fonction convexe est qu'un minimum local de la fonction est aussi un minimum global. Enfin, la production d'entropie est localement positive, comme précédemment. Ainsi, le principe combinant l'inégalité entropique, soit la production d'entropie positive, ainsi que le critère de convexité permet d'établir un système d'équations de champ de forme hyperbolique symétrique, car l'inégalité permet de restreindre la classe générique de champs.

Les problèmes d'ordre épistémologique que rencontrent les différentes théories de la thermodynamique de non-équilibre ne sont pas radicalement différents que ceux rencontrés en thermostatique. Que les quantités définissant l'état thermodynamique soient modifiées au gré des interactions entre le système et son environnement caractérisées par des flux et des forces ne constitue en rien un élément pour revoir la stratégie explicative de la physique. Le principal problème qu'elles rencontrent concerne leur testabilité. En effet, les valeurs spécifiques des différentes substance constituant les systèmes dont le comportement est susceptible d'être testé sont pour la plupart inconnues. Toutefois, certaines hypothèses concernant l'entropie peuvent être remises en question, mais celles-ci s'appuient généralement sur les résultats de la thermodynamique classique, comme celle de l'additivité, par exemple. C'est un élément de perplexité de la thermodynamique de non-équilibre que l'entropie thermodynamique soit reconnue comme étant limitée au phénomène d'équilibre mais qu'elle soit néanmoins utilisée pour expliquer les phénomènes de non-équilibre. Procède-t-on ainsi faute de mieux ? On peut dire que le contexte de la définition de la fonction d'état, pour une situation d'équilibre, est moins général que le contexte d'application du concept d'entropie.

4.2.7 Stratégie explicative

Si le second principe stipule que certains processus sont impossibles, le principe entropique introduit une fonction d'état dont la valeur ne peut qu'augmenter pour des systèmes isolés. L'idée générale de la stratégie explicative du principe entropique de la thermodynamique est la suivante : un système dans un état ayant certaines caractéristiques définissant l'équilibre peut atteindre un autre état par un processus ayant certaines caractéristiques (réversible ou irréversible) si et seulement une propriété de ce système qu'est l'entropie *totale*, au cours de ce processus, augmente ou, à la limite, demeure inchangée.

Ainsi, un processus n'est possible que si ce processus est compatible avec le principe entropique, soit $dS_T \geq 0$.

En d'autres termes, certains états sont inaccessibles à partir d'un certain processus. Par exemple, certaines réactions chimiques cessent avant d'être complétées, c'est-à-dire avant d'atteindre un état d'équilibre spécifique, parce que cet état posséderait une entropie moindre que celle de l'état initial sans surcompensation. Cette stratégie explicative est donc similaire à celle impliquant l'énergie où un état est accessible par un processus à partir d'un autre état si et seulement si l'énergie totale demeure constante. Cette dernière stratégie est donc moins restrictive que celle impliquant l'entropie, qui impose des conditions aux processus, car parmi l'ensemble des processus respectant la conservation de l'énergie, seul un sous-ensemble de processus respecte le principe entropique. Autrement dit, un processus respectant le principe entropique respecte celui de la conservation de l'énergie, mais l'inverse n'est pas nécessairement vrai. Cette restriction est particulièrement évidente dans le second principe de la thermodynamique où, alors que le premier principe autorise une interconvertibilité complète de la chaleur et du travail, il est stipulé par le second principe qu'une quantité de chaleur ne peut être convertie complètement en travail.

Il y a donc une restriction imposée aux processus énergétiques, c'est-à-dire que parmi l'ensemble des processus conservant l'énergie, seuls ceux compatibles avec une augmentation de l'entropie totale sont possibles. Autrement dit, toutes les transformations d'énergie respectant le principe de sa conservation ne sont pas nécessairement permises. C'est pourquoi il y a aussi une *direction* imposée à ces processus énergétiques. Pour bien comprendre ce que cela signifie, revenons à la ST : alors qu'une quantité de travail peut être convertie complètement en chaleur, une quantité de chaleur ne peut être complètement convertie en travail. Cela implique que soient bien définis les concepts de travail et de chaleur mais aussi que le système soit bien défini. On a vu que la ST était équivalente, dans certaines conditions, au principe entropique, en ce sens qu'une conversion complète d'une quantité de chaleur en travail implique non seulement un mouvement perpétuel de second type mais aussi une diminution totale d'entropie. Ainsi, d'une quantité d'énergie donnée pouvant être transférée sous forme de chaleur ou de travail, si celle-ci est transférée sous forme de chaleur, alors la quantité potentielle de travail qu'elle contenait ne peut être recouvrée. En revanche, le processus inverse consistant en une conversion complète du travail en chaleur et de manière équivalente à une augmentation totale d'entropie demeure possible – d'où cette directionnalité. Il a été vu que cette directionnalité est à la base de

l'irréversibilité de l'entropie thermodynamique, au sens de non-IRT mais aussi au sens d'irrécupérabilité⁵⁴.

Comme propriété ou variable d'état, l'entropie a des valeurs fixes pour des états fixes. Comme indiqué précédemment, la différence d'entropie entre deux états d'équilibre est indépendante de la manière ou du processus ou du chemin pris pour passer d'un état à l'autre, que le processus soit réversible ou non. Et comme indiqué aussi précédemment, l'état d'un système thermodynamique est complètement déterminé si au moins deux variables intensives sont aussi déterminées. Donc, si deux variables intensives sont spécifiées alors la valeur de l'entropie de cet état est aussi déterminée. Les différences d'entropie entre deux états ($S(x_i) - S(x_j)$) peuvent être déterminées de deux façons. L'une d'elle découle bien sûr de la définition de l'entropie et consiste à intégrer $\delta Q/T$ selon un processus réversible entre ces deux états. Ainsi, même si un quelconque processus irréversible relie deux états, il est possible en principe de calculer la différence d'entropie entre ces deux états en imaginant un processus réversible qui les relierait. L'autre façon consiste à s'appuyer sur des relations établies entre l'entropie et d'autres propriétés du système thermodynamique, comme, par exemple, l'équation centrale ou fondamentale. Mais ces relations sont généralement peu conviviales pour un calcul analytique, « à la main », surtout dans le cas de systèmes complexes, et nécessitent un calcul ou simulation informatique. Les résultats sont alors consignés sous forme de tables donnant les valeurs de propriétés comme la température, l'enthalpie et bien sûr l'entropie pour des états fixes⁵⁵.

Définir une stratégie explicative revient bien entendu à caractériser l'explanans mais aussi le domaine des explananda. L'élément fondamental de la stratégie explicative du principe entropique, comme on l'a vu, est la description d'un cas limite qui demeure par définition une idéalisation (voir aussi le Chapitre 2). Ce cas limite permet de *définir* une fonction d'état à partir de la description d'un processus réversible pour un système donné. De sorte qu'à *la limite* le changement d'entropie à la suite d'un processus réversible est déterminé par l'intégration de la quantité $\delta Q/T$ représentant la somme des quantités de chaleur échangées par le système à une température donnée. Lorsque cette quantité est nulle, si le système n'échange pas de chaleur ou s'il suit un processus réversible *et* cyclique, il va de soi que l'entropie ne peut ni augmenter ni diminuer. Le système peut être isolé ou non de son environnement. Dans le premier cas, selon le principe entropique, l'entropie du système ne peut qu'augmenter, tandis que dans le second cas, l'entropie peut augmenter ou diminuer, mais l'entropie totale, incluant celle de l'environnement, augmente nécessairement. Mais autant la condition de

⁵⁴ La situation est plus compliquée pour les systèmes microscopiques (voir le prochain chapitre).

⁵⁵ À titre d'exemple, la table de l'eau à la pression de 0,1 MPa donne une entropie de 152,390 kJ/kmol·°K à 100°K.

réversibilité que celle de l'isolation constituent des idéalizations. Le principe entropique stipule que tout processus s'écartant de ce cas limite ne peut être qu'une augmentation de l'entropie totale supérieure à la quantité $\delta Q/T$. De plus, par hypothèse, un tel processus ne peut être qu'irréversible, au sens donné par la non-IRT et l'irrécupérabilité. Ainsi, les processus susceptibles d'être observés et dont la description peut éventuellement corroborée sont donc des processus qui ne sont pas des idéalizations et qui sont donc irréversibles.

La stratégie explicative du principe entropique peut prendre des allures différentes selon que le système est isolé ou non, mais demeure fondamentalement la même. Dans un premier temps, un système isolé est censé atteindre l'équilibre thermodynamique, qui est considéré comme une situation où il n'y a plus de changement, du moins dans les paramètres observables caractérisant le système. Plus spécifiquement, s'il n'y a pas de transfert de masse, l'équilibre thermique signifie, par définition, qu'il n'y a plus de transfert de chaleur et par conséquent il n'y a plus de variation d'entropie. Ainsi, l'entropie d'un système isolé augmente avant d'atteindre un *maximum*. Et comme l'entropie dans cette condition ne peut diminuer, le système demeure dans cet état d'entropie maximale correspondant à un état d'équilibre. De sorte qu'une fois l'équilibre atteint, le système y reste à moins qu'il y ait intervention extérieure et que par le fait même le système ne soit plus isolé. Le système passe donc d'un état où ses paramètres (température, pression, etc.) ne sont pas uniformes à un état où ses paramètres le sont. Ce faisant, le processus qu'il suit est caractérisé par l'asymétrie temporelle de l'accroissement de son entropie⁵⁶.

Une précision est ici importante, car il peut sembler absurde de parler de l'augmentation d'entropie d'un système isolé, lequel ne peut recevoir de chaleur par définition. Il est toutefois possible qu'un système composé de deux corps à des températures différentes soit aussi isolé et que le transfert de chaleur d'un corps à l'autre corresponde à un processus vers l'équilibre thermique et à une augmentation d'entropie. Par exemple, une tasse de café chaud dans une salle de classe se refroidit, de la chaleur à une température donnée passe de la tasse à la salle et l'entropie du système composé de la tasse et de la salle, qui est un système isolé, augmente. Le système en question fait alors partie d'un système isolé. On peut décrire cette situation avec le concept, décrit plus tôt, de production d'entropie ou entropie interne.

Il y a ainsi une correspondance, pour ne pas dire une identification, entre un état d'entropie maximale et un état d'équilibre thermique. En clair, l'un explique l'autre : la tendance à l'équilibre est expliquée par le principe stipulant l'augmentation d'entropie. Ainsi, la question de savoir pourquoi

⁵⁶ Bien entendu, trouver une explication plus fondamentale de ce type de phénomène, un explanans de cet explanans, demeure un problème majeur de l'analyse fondationnelle de la thermodynamique et de la mécanique statistique.

l'entropie devrait être maximale à l'équilibre trouve sa réponse dans le principe entropique comme quoi l'entropie d'un système isolé ne peut pas décroître, et si elle décroît alors le système n'est plus isolé et n'est donc plus à l'équilibre. Un système qui n'est pas isolé peut toutefois être dans un état stationnaire, échangeant matière et/ou énergie avec son environnement, ce qui correspond à un état de non-équilibre et nécessite par le fait même une discipline traitant du non-équilibre. Dans ce cas, la situation se complique. Ainsi, dans un second temps, la stratégie explicative du principe entropique a prétention de rendre compte du comportement des systèmes qui ne sont pas isolés de leur environnement. Toutefois, ces systèmes sont typiquement dans un état de non-équilibre, ce qui sort du cadre théorique établie par la thermodynamique *qua* thermostatique. Néanmoins (pour l'instant), hormis cette idée générale de direction des processus de transfert et de transformation de l'énergie ainsi que cette tendance des systèmes isolés vers un équilibre thermique, la stratégie est aussi employée et parfois généralisée dans des contextes explicatifs plus particuliers mais néanmoins très divers. Pour un système donné pouvant échanger de l'énergie avec son environnement, l'entropie peut augmenter ou diminuer, mais si elle diminue pour le système ou pour l'environnement, alors elle *doit* augmenter d'une quantité supérieure pour le système ou pour l'environnement, respectivement. Comme on l'a vu, c'est cette idée de « surcompensation de l'entropie » qui permet, en principe du moins, au principe entropique de prétendre pouvoir s'appliquer aux systèmes qui ne sont pas à l'équilibre. Mais la possibilité d'une généralisation est soumise à la possibilité de « traduire » les divers échanges de matière et d'énergie d'un système et de son environnement en termes d'entropie, mais aussi et surtout à celle de la validation du concept même d'entropie dans pareils cas. Ces possibilités sont loin d'être évidente.

Mais voyons comment se déploie cette stratégie explicative dans ce type de situation. La stratégie n'est pas fondamentalement différente dans le cas du non-équilibre. Il faut seulement y ajouter les contraintes de détermination des différentes quantités caractérisant l'état d'un système changeant, et ce souvent rapidement, ainsi que les divers échanges avec son environnement. Mais la validité du principe est alors tributaire de la spécification du système impliqué. Car si le système n'est pas préalablement spécifié, clairement défini, il demeure possible d'identifier une région de l'espace-temps où il y a diminution d'entropie, une diminution locale, et d'en inférer – erronément – que le principe entropique ne tient pas. Dans ce cas précis, ce n'est pas la stratégie explicative du principe entropique qui est remise en cause mais plutôt que des limitations épistémiques posent les contraintes de son application à certains processus. Il y a toute une marge entre une difficulté pratique et une impossibilité de principe. C'est une raison importante soutenant la possibilité de généraliser la stratégie explicative du principe entropique.

4.2.7.1 STRATÉGIE GÉNÉRALISÉE

L'idée derrière cette généralisation de la stratégie explicative du principe entropique, que l'on pourrait nommer « stratégie généralisée », est la suivante : (i) toute transformation implique des échanges d'énergie et tout échange d'énergie implique une augmentation totale d'entropie, soit une diminution de la capacité d'un système isolé à effectuer un travail ; (ii) cette capacité est interprétée comme la possibilité d'obtenir un « niveau d'ordre », une « complexité organisée » (voir aussi Chapitre 5) ; (iii) enfin, étant donné cette tendance, cette direction des processus isolés vers l'augmentation d'entropie, cette capacité du système peut être maintenue par « surcompensation » dans un système ouvert, c'est-à-dire par une interaction avec son environnement de manière à diminuer l'entropie du système et à augmenter celle de l'environnement. Afin d'illustrer cette stratégie, prenons l'exemple d'un assemblage de briques : effectuer un travail sur ce système ne change pas son entropie, à moins que la friction dégage une quantité de chaleur, mais elle implique une augmentation d'entropie dans l'environnement ; si ce travail est utilisé selon l'information quant à la position des briques, à l'instar du bibliothécaire plaçant un livre sur un rayon en fonction de l'information indiquée sur le livre et sur le rayon, il est possible d'exploiter cette information afin de placer ces briques selon un certain ordre, plus grand qu'à l'état initial, et ce « degré d'ordre » peut être associé, du moins on le suppose ici, comme une diminution d'entropie ; on peut alors supposer que l'augmentation d'entropie dans l'environnement « surcompense » la diminution d'entropie dans le système. Cette « stratégie généralisée » est sujette à caution, mais elle exprime les possibilités explicatives du principe entropique, qui s'offre ici le concours d'une interprétation de l'entropie statistique.

Cette interprétation (i) de la capacité d'un système à effectuer un travail est bien entendu celle que sous-tend, comme vu précédemment, le second principe de la thermodynamique. Mais la corrélation inversement proportionnelle entre l'entropie et le travail n'est pas aussi assurée qu'on voudrait bien le croire, notamment dans les réactions chimiques. Une généralisation au-delà des systèmes mécaniques macroscopiques doit donc faire face à ce problème et la solution généralement adoptée consiste à emprunter la définition statistique de l'entropie, alors que le lien qu'elle peut avoir avec le concept de travail est plutôt ténu. Tel qu'indiqué, la partie (ii) de la stratégie fait appel à une interprétation particulière de la version statistique de l'entropie. Elle ne va pas de soi, mais elle est largement employée dans la littérature (Chapitre 5). Le troisième énoncé (iii) découle directement de la définition du principe entropique. Mais, encore une fois, il n'est pas évident que le concept d'entropie soit valide dans le cas des systèmes qui ne sont pas à l'équilibre, ou même qu'il le soit dans le cas des systèmes qui ne sont pas isolés de l'environnement où la déviation de l'équilibre serait considérée négligeable.

Cette généralisation impose donc que soient bien définies les variables à l'œuvre dans le second principe, soit la chaleur et le travail, mais aussi les limites du système ainsi que ces interactions avec l'environnement aussi sous forme de chaleur et de travail. Lorsque ces interactions sont violentes ou qu'elles impliquent des échanges de matière et des réactions chimiques complexes, cette interprétation devient plus ambiguë, voire inapplicable. Selon ce qui a déjà été dit à propos du second principe, ces interactions peuvent toutefois alimenter le système afin de maintenir, voire augmenter sa capacité à effectuer un travail. Par exemple, une voiture effectue des transformations d'énergie en consommant de l'essence, qu'elle transforme en énergie cinétique, au prix d'un rejet de chaleur et de matière dans l'environnement ; cette capacité à effectuer un travail va éventuellement diminuer, mais des interactions avec l'environnement, par l'ajout d'essence dans le réservoir ou même par des réparations, permettent de maintenir cette capacité à transformer l'énergie chimique en énergie cinétique. Le point crucial est d'associer certaines quantités échangées entre le système et son environnement à l'entropie ; par exemple, parmi ces quantités se trouve la chaleur dégagée par le moteur, ce qui est trivial en raison de la définition de l'entropie présentée plus tôt, mais aussi la matière comme des gaz d'échappement ou même les traces de caoutchouc laissées par les pneus de la voiture.

L'idée derrière cette généralisation consiste en ce qu'un système subissant des transformations d'énergie voit son entropie augmenter à moins d'entrer en interaction avec son environnement de manière à diminuer son entropie. Une telle généralisation joue souvent sur le plan de la métaphore où la matière et/ou l'énergie pouvant être utilisées, exploitées par le système afin de construire, d'augmenter le « niveau d'ordre » (Chapitre 5), sont associées à une capacité d'effectuer un travail (ce qui est plutôt naturel), mais où la matière et/ou l'énergie ne pouvant être exploitées par le système, généralement sous forme de chaleur, bref des « déchets », sont considérées comme de l'entropie. Sans une interaction particulière entre le système et son environnement certains états sont inaccessibles au système et *cet* état de fait est attribué à une augmentation d'entropie. La capacité d'effectuer un travail ou encore de maintenir un « niveau d'ordre » est représentée par la capacité de maintenir une certaine vitesse dans l'exemple de la voiture, mais elle est appliquée dans plusieurs autres cas, le plus évident étant celui du maintien de la vie où l'organisme doit tirer de son environnement les ressources nécessaires pour son métabolisme (Schneider & Kay 1994 ; Bejan 1998 ; Slathe 2004).

L'idée est générale et séduisante. Mais sa testabilité demeure difficile. L'appui sans doute le plus important à cette stratégie généralisée est le lien qu'elle présente avec l'impossibilité du mouvement perpétuel, mais il est loin d'être évident qu'elle puisse rendre compte des conditions à la définition de l'entropie. Afin de rendre compte de cette diminution de la capacité d'un système à effectuer un travail, l'entropie doit être associée à plusieurs états où la définition de l'entropie n'est pas facile. Dans l'exemple précédent de la voiture, la transformation de l'énergie chimique par consommation d'essence

entraîne tout un réseau de transformations énergétiques où l'on peut raisonnablement penser qu'il y a aussi changement d'entropie – la question est alors où, comment et de quelle valeur ? Ces transformations énergétiques sont en effet de plusieurs ordres, comme la friction des pistons du moteur impliquant un dégagement de chaleur, de sorte que l'augmentation d'entropie est assez bien définie. Mais en d'autres cas, comme celui de la combustion même du carburant ou celui des traces de pneus laissées par le détachement du caoutchouc, l'augmentation d'entropie est plus ou moins bien déterminée. Malgré les difficultés à déterminer les variations d'entropie, le second principe dans sa formulation ST et donc le principe quant à l'impossibilité du mouvement perpétuel de second type sont respectés dans l'exemple précédent, et même dans tous les cas possibles et susceptibles d'être corroborés, c'est-à-dire dans tous les cas où le principe de conservation de l'énergie est respecté et où les échanges de chaleur et de travail sont bien définis. Mais appliquer un principe comme celui de l'impossibilité du mouvement perpétuel aux systèmes complexes n'éclaire guère sur le comportement et les mécanismes de ces systèmes.

4.2.7.2 DISCUSSION : LIMITES ET CRITIQUES

L'unification d'un plus grand nombre de phénomènes par un même concept – l'entropie – peut être vue comme un signe de « progrès scientifique », au sens où il y a un certain succès explicatif (minimalement une dérivation des phénomènes à partir de la théorie) mais aussi un succès prédictif (Kitcher 1989 ; Kuipers 2007). Cette affirmation se base alors sur des valeurs méthodologiques comme l'unité, l'efficacité ou la simplicité du discours scientifique. Mais il n'est pas exclu que ces valeurs s'appuient sur un arrière-plan métaphysique de tradition parméniennienne traduisant l'uniformité, l'homogénéité et la simplicité de la nature. Or ce progrès n'est réel que si l'unification est minimalement effective, c'est-à-dire si le concept soi-disant unificateur est applicable et bien défini pour les phénomènes ainsi prétendument unifiés dans leur description, mais aussi, de manière plus pertinente, au sein d'une même stratégie explicative. Il ne faut pas, en effet, qu'il y ait une application triviale à des non-exemples, comme des faux positifs, à l'instar d'une théorie à propos des corbeaux noirs « corroborée » par des non-corbeaux blancs. En tentant d'expliquer le principe de l'impossibilité du mouvement perpétuel et le second principe de la thermodynamique, ou dans une moindre mesure de les exprimer dans un autre contexte théorique jugé valide et sans doute plus sûr, la stratégie explicative semble ainsi se restreindre à une plus petite classe de phénomènes : il n'est pas clair que les conditions auxquelles doit satisfaire le concept d'entropie soient moins restrictives que celles des concepts de chaleur et de travail. De fait, on reproche au concept d'entropie de reposer sur des idéalizations et il semble, quoique cela mérite examen, que la complexité des systèmes (possédant plusieurs inter-relations entre sous-systèmes) soit davantage problématique à son application que les

concepts de chaleur et de travail. Une interprétation plus libérale du concept d'entropie, laissant de côté l'exigence d'une détermination quantitative dans les cas plus ambigus, permet toutefois d'étendre cette classe de phénomènes au sein d'une stratégie généralisée, intuitive, mais théoriquement plus ambiguë.

En outre, que ce progrès, le cas échéant, soit un cas de réduction théorique réussie est aussi ambigu. Car l'une des conceptions d'une réduction exige que la théorie réductrice soit plus générale que la théorie réduite, en ce sens où elle doit rendre compte de plus de phénomènes observables. Il faut alors remarquer deux choses ici. D'abord, les rôles respectifs de la théorie réductrice et de la théorie réduite ne sont pas évidents puisque la thermodynamique dite classique contient déjà le concept d'entropie, et la théorie sans ce concept serait quelque chose comme la « thermodynamique préhistorique » similaire à la théorie de Carnot (sans le concept de calorique). Il y a alors un travail de reconstruction théorique par rapport à la théorie réduite car elle n'a jamais été élaborée à proprement parler. Celle-ci contiendrait sans aucun doute le principe de l'impossibilité du mouvement perpétuel et la notion d'équilibre thermique. Elle pourrait se rapprocher, par ailleurs, de la ST. Ensuite, une fois la théorie réduite bien identifiée, et si elle prend les aspects mentionnés, il faudrait déterminer si oui ou non la thermodynamique classique contenant le concept d'entropie permet de rendre compte de plus de phénomènes que cette « thermodynamique préhistorique ». Comme on pourrait s'y attendre, la réponse n'est pas aussi claire qu'on le souhaiterait. Car la première permet de rendre compte de phénomènes plus divers, comme certains phénomènes chimiques, mais elle se « limite » à la classe de phénomènes où l'entropie est bien définie et il est plausible que cette classe soit plus restreinte que celle où la chaleur et le travail sont bien définis.

La stratégie explicative du principe entropique, comme vu précédemment, permet néanmoins de rendre compte de l'irréversibilité des phénomènes thermodynamiques *et* de la tendance à l'équilibre des systèmes thermodynamiques. Mais la particularité de cette stratégie, déjà notée avec la théorie de Carnot et la ST, réside en ceci que les processus les mieux définis sont ceux relevant d'une idéalisation. Pour les processus réversibles et à l'équilibre, une quantité maximale de travail peut être obtenue, pour une différence de température donnée, et l'entropie produite est alors minimale. Mais ce cas n'est jamais instancié, il n'est qu'approché. Tous les processus dont la description théorique est susceptible d'être corroborée ne peuvent donc satisfaire ces conditions. Ils sont alors évalués par rapport à l'écart qu'ils présentent avec ce cas idéalisé. L'idée est alors que plus un processus s'éloigne du cas idéalisé, plus il produit de l'entropie et plus il est irréversible. Il a été vu que cette stratégie ne posait pas à prime abord de problème épistémologique majeur, mais un examen plus approfondi est de mise.

Il a été vu (Chapitre 2) que les idéalizations abondent en sciences. La question est de savoir si leur utilisation est justifiée, si le fait qu'elles n'offrent pas une description exhaustive, « littérale » de la réalité puisse être rédhibitoire. Ceci dit, il n'est pas à prime abord injustifié de faire appel à un principe stipulant ce qui est impossible car, au demeurant, il n'est pas nécessairement infalsifiable. De plus, certains énoncés scientifiques ne sont pas falsifiables bien qu'ils soient testables (par ex. il y a des positrons). Contrairement aux fictions, qui sont aussi fréquemment utilisées en sciences, les idéalizations approximeraient ainsi les cas réels et posséderaient donc un caractère de « vérissimilitude » (Chapitre 2). Il peut permettre entre autres des prédictions fort utiles et stimuler la compréhension des phénomènes. Mais ce qui importe ici est de noter que la stratégie explicative du principe entropique offre un paramètre d'évaluation des phénomènes dont la description théorique est susceptible d'être corroborés où l'« écart » d'avec le cas limite constitue la base de l'*explanans* : plus l'entropie est élevée, plus le comportement du système s'éloigne d'un processus réversible.

La particularité de cette stratégie, où l'entropie n'est clairement définie que dans ce cas limite d'un processus réversible, nous force donc à nous s'interroger sur le lien entre la définition d'un cas limite et celle d'une fonction d'état. Par contraste, en ce qui concerne la mécanique newtonienne une fois de plus, celle-ci définit le mouvement comme un état et non comme un processus, contrairement à la mécanique aristotélicienne, par exemple, où le mouvement est un « processus de changement » (Koyré 1966 : 201) vers un état de repos. Puisqu'une force est aussi considérée comme la cause d'un mouvement uniforme, tout corps qui n'est pas au repos, dans cet état « naturel », doit ainsi recevoir une explication. En revanche, la mécanique newtonienne conçoit le mouvement comme un état pouvant être décrit par une vitesse où le cas d'une vitesse nulle, soit un état de repos, n'est qu'un cas limite sans différence fondamentale avec celui d'une vitesse non nulle. Le cas d'un mouvement rectiligne uniforme est, par contre, un cas limite idéalisé. Le comportement de tout corps qui s'écarte de ce cas idéalisé, donc de tout corps susceptible d'être observé, peut être expliqué par l'action d'une « force déséquilibrée » (non-balancée), comme résultante d'une somme vectorielle, à l'origine de cet écart entre le mouvement observé et le cas limite. Or, seules les forces déséquilibrées ont des effets observables⁵⁷. Car une force déséquilibrée est la cause d'un changement de mouvement, comme une accélération ou un changement de direction (impliquant aussi une accélération), et seul ce dernier est observable. Éliminer ce cas limite idéalisé de la stratégie explicative revient à rendre inopérant le concept de force déséquilibrée dans l'explication d'un changement de mouvement. Contrairement à la

⁵⁷ Typiquement, une force équilibrée (balancée) n'a pas d'effet observable en dynamique. Par exemple, selon la troisième loi de la mécanique newtonienne, le poids d'une table immobile exerce une force sur le plancher qui exerce une force de même intensité, mais de signe opposé, sur la table. Mais il est vrai qu'une force équilibrée (balancée) puisse produire une déformation observable.

mécanique aristotélécienne, un corps en mouvement rectiligne uniforme n'a pas besoin d'être expliqué puisqu'il est défini par la loi de l'inertie elle-même, par la définition même des concepts de la première loi, et nul besoin de faire appel à une force ou à une cause pour expliquer ce mouvement idéalisé.

En définissant le mouvement rectiligne uniforme, la mécanique définit un cas limite où il n'y a aucune accélération, aucune variation de vitesse dans les trois directions spatiales. En corollaire, une accélération d'un système dans l'une ou l'autre de ces trois directions spatiales implique l'action d'une force, d'une « cause », dans la direction correspondante et éloignant le système du cas idéalisé de mouvement rectiligne uniforme. De manière similaire, en définissant un processus réversible, la thermodynamique définit un cas limite où il n'y a aucune irréversibilité, aucune production d'entropie excédentaire, ce qui a été appelé plus tôt « production irréversible d'entropie », de sorte que la différence d'entropie entre l'état initial et l'état finale correspond à la valeur de l'intégrale $\int \delta Q/T$ ⁵⁸. En corollaire, une production d'entropie excédentaire, une « production irréversible d'entropie », implique une irréversibilité, une « force », une « cause » en quelque sorte, éloignant le système du cas idéalisé de processus réversible. Typiquement, une telle « force » est une différence de température menant à un transfert irréversible de chaleur sans production de travail, lequel est donc accompagné d'une augmentation d'entropie. Ce qui semble effectif dans cette stratégie explicative est précisément l'action d'une « force » ou d'une « cause » faisant partie de l'*explanans* pour des situations s'écartant du cas limite, en bref pour *tous* les cas susceptibles d'être observés et dont la description théorique serait éventuellement corroborée. Que la description de certains phénomènes ne soit qu'une idéalisation jamais instancié ne pose pas de problème au sein de cette stratégie explicative puisqu'elle sert à définir l'*explanans*, dont le concept d'entropie lui-même. En effet, cette limite fixée par l'idéalisation permet de définir un écart entre cette même limite et le comportement de certains systèmes devant être expliqué. Par exemple, une « force » explique l'écart entre le mouvement d'un corps particulier et le mouvement rectiligne uniforme étant un cas limite, toute comme une « irréversibilité » explique l'écart entre le comportement d'un système particulier et un processus réversible aussi idéalisé.

Ainsi, au sein de la stratégie explicative du principe entropique, l'idéalisation d'un processus réversible ne pose pas plus de problème que celle du mouvement rectiligne uniforme au sein de la mécanique newtonienne. La présentation des étapes de la construction de la fonction d'état a révélée que cette idéalisation permettait justement de définir cette fonction d'état lorsque les « valeurs d'équivalence », selon l'expression de Clausius, se compensent mutuellement. Un type de problème qui peut être soulevé est celui, déjà mentionné, à l'effet que les systèmes ne se comportent *pas*, du point de

⁵⁸ Contrairement à une situation de processus irréversible où la différence d'entropie entre l'état initial et l'état final est *supérieure* à la valeur de l'intégrale $\int \delta Q/T$.

vue de la physique, tels que les décrivent les exemples employés dans l'argumentation visant à établir le principe entropique. Mais le point crucial de l'argumentation est justement de remarquer qu'il s'agit d'une idéalisation, d'un cas limite. C'est-à-dire qu'il faut abstraire d'une description aussi fidèle de la réalité que faire se peut ce qui est, justement, susceptible de « causer », de provoquer un « écart » d'avec ce cas limite idéalisé. L'identification de ce qui « cause » ou provoque un « écart » d'avec ce cas limite idéalisé, comme une différence de température provoquant un transfert d'une quantité de chaleur supérieure à la quantité minimale établie par ce cas limite, lequel correspond à un processus réversible comme un cycle de Carnot, un transfert qui s'effectue vers l'environnement, permet de former une partie essentielle d'un *explanans* de type causal.

D'aucuns peuvent soutenir que les systèmes ouverts ne constituent pas des systèmes thermodynamiques puisqu'ils ne sont pas à l'équilibre, et qu'en conséquence, il serait vain de leur attribuer une quelconque valeur d'entropie. Cependant, c'est justement par une interaction d'un système avec son environnement, ce qui en fait un système ouvert, qu'il est possible de préparer des états de faible entropie. Autrement, un état de faible entropie ne serait jamais observé, car le principe entropique stipule que l'entropie des systèmes isolés ne peut qu'augmenter. Bien que cette intervention ne soit pas caractérisée clairement par la théorie, elle en constitue un élément implicite essentiel. Mais la préparation d'un état de faible entropie exige un certain délai permettant au système de retrouver un état d'équilibre après l'intervention. Donc, le principe entropique permet d'identifier les systèmes qui ne sont pas isolés si leur entropie diminue et stipule que la seule façon de faire diminuer l'entropie d'un système est par une intervention avec son environnement, c'est-à-dire en étant un système ouvert. Mais l'inverse n'est pas vrai, car ce n'est pas parce que le système est ouvert que nécessairement son entropie diminue. La distinction entre système isolé et système ouvert est liée de près à celle de processus « spontané » et processus « non-spontané » ou « forcé » (déjà présenté), en ceci qu'un processus « spontané » est défini comme un processus se produisant en l'absence d'apport d'énergie de l'environnement d'un système suivant ce processus, bien qu'il puisse transférer une quantité d'entropie dans cet environnement. Ainsi, un processus spontané implique une augmentation de l'entropie totale mais peut aussi impliquer une diminution locale d'entropie. De sorte que l'explication d'un processus spontané où il n'y a pas d'apport d'énergie de l'environnement peut se rapporter au principe entropique autorisant ce type de processus s'il y a augmentation de l'entropie totale.

Un dernier point avant de conclure. La ST exprime clairement que d'une différence de température un flux de chaleur et une quantité de travail peuvent être obtenus. La disparition d'une différence de température entraîne par conséquent la disparition du flux de chaleur et de la quantité de travail correspondante. Étant donné la tendance, *ceteris paribus*, des systèmes à l'équilibre thermique, donc à la disparition des différences ou gradients de température, on peut inférer la tendance à la perte de

capacité des systèmes à effectuer un travail. Si cette tendance est généralisée à l'univers au grand complet, alors on peut inférer cette « mort thermique » correspondant à un état où il n'y a pas de transfert chaleur ni donc de conversion de chaleur en travail, ni variation d'entropie. Comme il faudrait une quantité de travail pour reproduire un gradient de température et que cette quantité ne peut provenir d'une conversion de la chaleur en travail dans une situation d'équilibre, on peut se poser la question si elle est seulement possible et quelles en sont les conséquences. Partant, on peut aussi se poser la question de savoir si un monde dans un état de « mort thermique » est compatible avec des systèmes échangeant de l'énergie. Or un monde dans un tel état ne semble pas compatible avec la vie, car tout être vivant échange matière et énergie avec son environnement, comme quoi la vie semble impliquer un état de non-équilibre (Sheehan 2007). Il s'agit sans doute de l'idée derrière l'affirmation de Kelvin (1852) sur le destin de l'humanité, comme quoi la terre deviendra « inadéquate à l'habitation de l'homme tel qu'il est présentement constitué ». Le moins que l'on puisse dire est que les conséquences eschatologiques de cette affirmation s'appuyant sur le second principe pour inférer la disparition de la vie contraste *prima facie* avec celles voulant que la vie elle-même soit une « manifestation du second principe » (Schneider & Kay 1994), ou que le principe entropique soit l'outil de prédilection pour comprendre l'évolution de la vie en général (Brooks & Wiley 1988) ou des écosystèmes (Ludovisi, Pandolfi & Taticchi 2005), ou encore la sélection naturelle (Kaila & Annala 2008 ; Rosenberg 2011). La richesse et la complexité du principe entropique se dévoile ici dans toute sa splendeur (Chapitre 6)⁵⁹.

4.2.7.3 REMARQUES CONCLUSIVES

Il est donc possible d'expliquer le comportement des systèmes thermodynamique en général, et plus particulièrement leur irréversibilité (non-IRT et irrécupérabilité), de manière déductivo-nomologique en subsumant ce qui doit être expliqué (*explanandum*) sous ce qui explique (*explanans*), c'est-à-dire en déduisant (rétrospectivement) un phénomène particulier de la conjonction d'énoncés généraux comme le principe entropique et de conditions initiales. L'*explanans* inclut alors un cas limite idéalisé par rapport auquel le comportement des systèmes observés est évalué. Ce cas limite idéalisé correspond à un processus réversible où le changement d'entropie est défini par une fonction d'état telle que $dS = \int \delta Q/T = S(x_f) - S(x_i)$. Ce faisant, l'*explanans* inclut aussi des « causes » et prendre la forme d'une

⁵⁹ À titre d'exemple de la diversité des rôles que l'on veut faire jouer au concept d'entropie, en terminant leur chapitre consacré à ce concept, Van Wylen, Sonntag & Desrochers (1992 : 253) s'interrogent sur « l'avenir du monde naturel tel que nous le connaissons » et si « toutes les évolutions que nous connaissons s'accompagnent d'un accroissement de l'entropie ». Sans y répondre, pour des raisons évidentes, ils notent cependant : « nous pouvons considérer le second principe de la thermodynamique comme une façon d'expliquer l'œuvre antérieure et continue d'un créateur qui détient aussi la réponse à notre destinée future et à celle de l'univers » !

explication causale dans son acceptation la plus générale. En effet, des irréversibilités (qui ne sont pas toujours bien définies) sont considérées comme des « causes » (au demeurant un terme vague) du comportement du système s'écartant de celui du cas limite. La stratégie explicative est donc double, selon la classification des théories de l'explication, en ceci qu'elle adopte une approche inférentielle *et* causale, à l'instar de la mécanique classique, et plus particulièrement sa première loi. Mais il est vrai que plusieurs théories causales de l'explication incluent les lois dans leurs modèles afin de caractériser les causes.

La stratégie explicative est souvent généralisée à des systèmes complexes où divers transferts de matière et de chaleur sont associés à des transferts d'entropie. Ces derniers sont à leur tour identifiés à une diminution de la capacité du système à effectuer un travail, c'est-à-dire qu'un système qui reçoit une quantité d'entropie (plutôt : certains transferts de matière et de chaleur) subit cette diminution et vice-versa, ou encore, selon une interprétation importée de la version statistique de l'entropie, à une augmentation du « désordre », c'est-à-dire qu'un système qui rejette une certaine quantité d'entropie augmente son niveau d'organisation et vice-versa. Si l'association des transferts de matière et de chaleur à des transferts d'entropie pour des systèmes complexes et/ou loin de l'équilibre de même que l'interprétation du « désordre » sont justifiées, alors cette stratégie explicative généralisée peut être très puissante puisqu'elle peut rendre compte d'une très vaste gamme de phénomènes. Bien que cette « importation » conceptuelle, de l'hydrodynamique dans le premier cas ou de la mécanique dans le second cas, soit discutable, le principe entropique peut s'appuyer sur la « robustesse » du principe de l'impossibilité du mouvement perpétuel.

4.3 Conclusion

Le concept d'entropie défini par la fonction d'état qui porte son nom a de grandes ambitions. Celles-ci s'appuient sur l'argument qu'il exprimerait le second principe de la thermodynamique et plus spécifiquement l'impossibilité du mouvement perpétuel de second type. Mais cet argument a ses failles. Le caractère fondamental du second principe de la thermodynamique se situe dans la conversion de la chaleur en travail. Le concept d'entropie thermodynamique trouve sa pertinence principalement dans sa capacité de rendre compte des changements d'états des systèmes opérant de telles conversions. Pour être en mesure de le faire la température des systèmes en jeu doit aussi être bien définie. Or, certains systèmes reçoivent difficilement une application claire des concepts de chaleur, de travail et de température. C'est le cas notamment des systèmes quantiques ou des systèmes loin de l'équilibre.

Les prétentions du concept d'entropie ont été grandes, et ce particulièrement en ce qui concerne l'explication des asymétries temporelles, de l'irréversibilité. L'inexorable augmentation d'entropie serait à la « base » de plusieurs asymétries tout ce qu'il y a de plus intuitives, comme celle, notamment, entre la connaissance du passé et la possibilité de changer le futur, et celle entre une cause et son effet. Mais à l'inverse on peut soutenir que cette augmentation d'entropie serait due à des propriétés fondamentales du temps. Pour rendre compte de l'une de ces deux thèses il faut revenir à la question du temps lui-même et cerner correctement l'argumentation visant à définir le temps à partir de la description des phénomènes, ou inversement. Il faut donc, d'abord, distinguer l'asymétrie des processus avec celle du temps, ce qui nécessite des repères conceptuels, au premier chef le concept d'irréversibilité. Celui-ci reçoit principalement deux significations : l'irréversibilité au sens de non-invariance par renversement du temps (non-IRT), la séquence des états ne pouvant être inversée, et l'irréversibilité au sens d'irré récupérabilité, soit la non-disponibilité d'un processus pouvant récupérer l'état initial sans modifier l'environnement. Que certains processus soient linéaires ou cycliques n'implique pas que le temps le soit, car tant qu'un processus irréversible comme un enregistrement demeure possible, une histoire du monde est possible au sens où chaque événement du monde, entendu comme l'univers en entier, est unique, même si certains processus locaux peuvent se répéter. Donc, la thèse voulant que l'entropie puisse être à la « base » de l'asymétrie du temps lui-même doit pouvoir ramener ce concept à l'univers en entier, à l'ensemble des systèmes du continuum spatio-temporel, pour « tout temps ». Mais la thermodynamique ayant donné naissance au concept d'entropie n'a pas une prétention aussi universelle.

Dans une analyse fondationnelle de la physique, il importe de distinguer le second principe de la thermodynamique du principe entropique. Bien que la littérature procède généralement à une

identification de ces principes et qu'ils peuvent mener à des conclusions similaires dans certains contextes, leur dérivation ne repose pas sur les mêmes hypothèses, ils n'ont pas la même signification et leur interprétation n'ont donc pas les mêmes conséquences épistémologiques.

En termes simples, le second principe de la thermodynamique vient mettre un terme à ce « désir pathétique » d'obtenir quelque chose à partir de rien. La signification du second principe a été dégagée ici par une reconstruction théorique basée sur quelques principes et appelée *séquence thermodynamique* (ST), qui stipule en substance qu'une quantité de chaleur ne peut être complètement convertie en travail. En ce sens, il décrit l'impossibilité d'un mouvement perpétuel de second type, soit un processus permettant de convertir toute la chaleur d'un corps en travail. En conséquence, dans la séquence inverse, un transfert de chaleur du froid vers le chaud est impossible sans dépense de travail. La robustesse de ce principe exprimant cette impossibilité n'est plus remise en doute, mais, principalement, pour des raisons de fait et non de principe. Il a été vu que le caractère « négatif » de ce type d'énoncé ainsi que l'expression d'un contrefactuel ne posaient pas de problème épistémologique majeur. De plus, un processus thermodynamique est non-invariant par renversement du temps (non-IRT) car tout processus ne peut satisfaire ce cas limite d'un rendement maximal. Un processus thermodynamique est aussi irrécupérable car l'unique façon de récupérer un état initial est d'ouvrir le système et d'échanger de l'énergie et éventuellement de la matière avec l'environnement, lequel ne peut par conséquent demeurer inchangé, ce qui ne peut satisfaire le critère d'irréversibilité au sens d'irrécupérabilité.

Ce cas limite est défini à partir de l'hypothèse d'un processus cyclique réversible où un système échange des « valeurs d'équivalence », tel que $\sum f(T_i)Q_i = 0$ ou, dans le cas continu, $\oint f(T)Q = 0$. En déterminant la fonction $f(T_i)$ comme étant $1/T$, et puisque l'équation $\oint f(T)Q = 0$ est indépendante du chemin d'intégration, du processus reliant les états initial x_i et final x_f , alors on peut définir une fonction d'état telle que $S_T \equiv \int dS = S(x_f) - S(x_i) = \int \partial Q/T$, donc $S(x_f) = S(x_i) + \int \partial Q/T$. Lorsque la condition d'un processus réversible n'est pas satisfaite, $\int dS > \int \partial Q/T$, de sorte que $S(x_f) > S(x_i) + \int \partial Q/T$. Ainsi, l'entropie n'est pas une propriété conservative et il n'y a donc pas une telle chose que la loi de la conservation de l'entropie. Aussi, la fonction S_T est considérée comme additive et extensive, en ce sens que l'entropie d'un système composé de plusieurs parties correspond à la somme des entropies des parties. Enfin, puisque l'entropie thermodynamique est une variable définie par une fonction d'état, le changement d'entropie d'une substance quelconque passant d'un état à un autre est la même pour *tout* processus, qu'il soit réversible ou irréversible. En termes clairs, si l'entropie d'un état x_i est $S(x_i)$ et celle d'un état x_f est $S(x_f)$, alors la différence d'entropie $S(x_i) - S(x_f)$ est la même peu importe le processus reliant les états x_i et x_f .

La fonction d'état définissant l'entropie permet ainsi une traduction du second principe de la thermodynamique vers le *principe entropique*. Ce principe stipule que la variation *totale* d'entropie ne peut qu'augmenter. La variation totale d'entropie, dS , peut être scindée en deux, soit en une production d'entropie ou entropie interne, $d_i S$, et un *flux* d'entropie ou entropie externe, $d_e S$. Ainsi, l'inégalité $\int dS = S(x_f) - S(x_i) \geq \int \partial Q/T$ rappelle que la production d'entropie est toujours supérieure au flux d'entropie ; ou encore, le changement d'entropie *total* est toujours supérieur ou égal à 0. Dans le cas d'un système isolé, $d_e S = 0$ par définition, et dans le cas d'un système fermé ou ouvert, $d_e S$ demeure indéterminée et dépend des conditions du système et de son environnement, de leur interaction. Aussi, de par les propriétés de la fonction S établies précédemment, pour un processus réversible $d_i S = 0$ et pour un processus irréversible $d_i S > 0$, de sorte que pour tout processus $d_i S \geq 0$. En conséquence, dS est nulle si et seulement si le système est isolé et subit un processus réversible. Le principe entropique peut donc se résumer ainsi : $dS \geq 0$, ce qui équivaut à $S(x_f) \geq S(x_i)$.

La particularité de la stratégie explicative du principe entropique réside dans la description d'un cas limite idéalisé à partir duquel les systèmes susceptibles d'être expliqués sont évalués. Dans ce cas limite, $\int dS = \int \partial Q/T$ et dans *tous* les autres cas $\int dS > \int \partial Q/T$. Le point crucial de cette stratégie est donc que la fonction d'état est *définie* à partir d'une argumentation faisant appel à des idéalizations et que les explananda ne peuvent être décrit, par voie de conséquence triviale, par les mêmes « outils » de cette argumentation puisque, justement, celle-ci fait appel à des idéalizations. Il n'est donc pas étonnant que cette stratégie explicative, en se référant à ce concept qu'est l'entropie dont la définition repose sur certaines idéalizations, fasse référence à des systèmes qui ne sont pas « réels » et qu'elle fasse des énoncés « négatifs » ou contrefactuels, si tant est que les idéalizations ne décrivent pas ce qui est. La stratégie explicative est donc double en ceci qu'elle adopte une approche inférentielle *et* causale, à l'instar de la mécanique classique, et plus particulièrement sa première loi, où une « cause » *explique* l'écart du comportement décrit dans l'explanandum avec celui du cas limite nomologiquement idéalisé. Enfin, la stratégie explicative est souvent généralisée à des systèmes complexes (Chapitre 6) où divers transferts de matière et de chaleur sont associés à des transferts d'entropie, qui sont à leur tour identifiés à une diminution de la capacité du système à effectuer un travail ou, selon une interprétation particulière qu'il faut justifiée, à une augmentation du « désordre » (Chapitre 5).

5 Critique du concept d'entropie II

RÉSUMÉ. Le concept d'entropie statistique est censé rendre compte du concept d'entropie thermodynamique. Cette affirmation nécessite déjà des précisions puisque le rapport épistémique (explicatif, réductif, interprétatif, etc.) entre ces deux concepts n'est pas clair, d'autant plus que ces concepts sont polysémiques. Néanmoins, deux choses semblent assurées : le concept d'entropie statistique établit un « pont » entre les descriptions macroscopiques de la thermodynamique et celles microscopiques de la mécanique classique, et ce faisant il fait intervenir les probabilités. Plusieurs interprétations ont été proposées, comme celles des probabilités, de l'information manquante, de la perte de travail et de la dispersion. À terme, on verra que l'entropie statistique est un concept physique qui ne reçoit une connotation subjectiviste qu'en raison d'une certaine interprétation des concepts mêmes qui sont mobilisés pour l'interpréter. Il appert donc qu'il a tout ce qu'il faut pour offrir un espace de discussion philosophiquement riche.

Never mistake motion for action.

Ernest Hemingway

Un professeur qui demandait à chaque semestre la même question à l'examen final, à savoir : « qu'est-ce que l'entropie ? », se vit suggérer par son assistant de modifier la question afin de placer les étudiants dans l'expectative ou à tout le moins de créer un peu de nouveauté. À cela le professeur répondit : « c'est toujours la même question, mais la réponse change à chaque fois ». Cette anecdote, qu'elle soit véridique ou non, illustre bien la polysémie qui traverse le concept d'entropie, mais aussi une certaine acceptation de cette situation singulière, une sorte de résignation. Pour le meilleur et pour le pire, on semble s'en accommoder¹.

Dans le chapitre précédent, l'entropie thermodynamique a été présentée comme une fonction d'état tirée d'un cas limite idéalisé. Le second principe de la thermodynamique, stipulant l'impossibilité d'une conversion complète de la chaleur en travail, soit l'impossibilité du mouvement perpétuel de second type, a été distingué du principe entropique, qui stipule en substance que l'entropie totale d'un système et de son environnement ne peut qu'augmenter. En ce sens, un processus réversible pour lequel l'entropie est définie constitue un cas limite jamais instancié. Cette irréversibilité a donc pu être une

¹ « *For completeness, you may add your own personal favorite here:* » ; Capek & Sheehan (2005 : 23), présentant une liste (non exhaustive !) des différentes formulations de l'entropie.

motivation d'une reformulation, d'une traduction sur d'autres bases théoriques face à la réversibilité de principe des lois physiques, au premier chef celles de la mécanique classique.

C'est précisément au sein de cette discipline que le travail de reformulation s'est opéré. Il a été vu (Chapitre 3) que les travaux de pionniers comme Clausius, Maxwell et Boltzmann tentaient d'expliquer les systèmes typiquement thermodynamiques par des procédures statistiques appliquées à de grands ensembles de particules. Mais les paramètres dits macroscopiques, puisqu'ils sont observables ou mesurables, ne spécifient pas l'état dit microscopique déterminé par l'ensemble des paramètres dynamiques pour chacune des particules. Ainsi, plusieurs états microscopiques sont compatibles avec un seul état macroscopique. La stratégie orthodoxe de formalisation en mécanique statistique consiste à ainsi définir une fonction d'une dynamique sur l'espace des phases canonique. Autrement dit, elle repose sur l'idée que le comportement des systèmes macroscopiques ou phénoménologiques, bref thermodynamiques, dépend d'une manière ou d'une autre sur la dynamique microscopique de leurs constituants que sont les particules, les molécules. Cependant, cette dynamique n'est pas tractable pour les systèmes d'une grandeur suffisante pour être justement considérés comme macroscopiques et c'est pourquoi l'approche statistique est adoptée. C'est-à-dire qu'il ne s'agit pas de retracer chaque molécule afin d'en décrire l'histoire mais plutôt d'en déterminer le comportement global en calculant des valeurs statistiques comme des moyennes. Cette méthode requiert des hypothèses simplificatrices qui posent des problèmes conceptuels d'envergure quant à la validité des résultats de cette approche.

Or, l'entropie statistique est en quelque sorte une quantification de cette compatibilité. Il n'est donc pas surprenant, compte tenu de ce rapport « un-plusieurs » entre un état macroscopique et des états microscopiques, que l'entropie ait été associée à la probabilité de cet état macroscopique. Même, compte tenu de l'augmentation d'entropie stipulée par le principe entropique, certains ont avancé, et Boltzmann lui-même, que les systèmes thermodynamiques passaient d'états peu probables à des états plus probables. C'est pourquoi ce chapitre s'ouvre sur une importante discussion sur les probabilités. Non pas sur le calcul des probabilités, ou sur leur axiomatisation, mais plutôt sur leur signification et leur interprétation en physique.

Vient ensuite une présentation de quatre versions différentes de l'entropie en mécanique statistique, celles de Boltzmann et celles de Gibbs. Dans certains contextes, ces différentes définitions peuvent être équivalentes, mais elles présentent toutes leurs propres « idiosyncrasies » et leurs propres conséquences interprétatives. Le problème est qu'il y a polysémie sans qu'il n'y ait de description des domaines de validité respectifs correspondant aux multiples définitions et interprétations, ce qui se ramène en fait à la signification même d'une polysémie, car, dans le cas contraire où il y aurait plusieurs sens *avec* circonscription des domaines de validité respectifs, alors ce serait plutôt une relation de synonymie ou

encore des concepts ni plus ni moins différents. C'est pourquoi on peut bien soutenir qu'il n'y a « pas de définition complètement satisfaisante de l'entropie » (Càpek & Sheehan 2005 : 23).

Cinq interprétations différentes de ce concept sont discutées (section 5.3.4). Malheureusement (ou heureusement), chacune de ces quatre interprétations apporte un éclairage pertinent sur le concept d'entropie ou plus précisément, leur examen critique soulève des questions pertinentes à son élucidation. Bien évidemment, interpréter l'entropie avec le concept de probabilité est inévitable, mais il est très discutable de justifier l'augmentation d'entropie dans le temps par le calcul des probabilités. Cette sous-détermination du micro-système est responsable de bien des problèmes épistémologiques, comme la validité des hypothèses sur les inobservables ou l'interprétation des descriptions formelles de rapport « un-plusieurs », de la réalisabilité multiple (Chapitre 2). Celles du désordre et du manque d'information en sont des exemples.

Le concept d'entropie est donc rempli de promesses mais reste à main égale ambigu. D'où l'intérêt d'un examen critique. Il est en outre associé à des concepts d'intérêts philosophiques, comme l'irréversibilité, la causalité, la probabilité et l'accessibilité. Mais il intervient aussi au sein d'explications scientifiques particulières comme celles décrivant un processus vers un état d'équilibre, ou celles définissant la probabilité d'un état microscopique à partir d'un état macroscopique.

5.1 Probabilités

5.1.1 Introduction

Les probabilités sont partout : elles sont utilisées en sciences naturelles et en sciences humaines, en gestion de risques, la météo nous les présente afin de planifier notre journée, l'éthique les utilise pour orienter le bon choix parmi les possibles, et la philosophie des sciences tente de les interpréter. Cette diversité d'usage du concept de probabilité, issue en large partie de la diversité des disciplines qui l'emploient, crée aussi une diversité de signification et d'interprétation. Ce concept a aussi subi plusieurs modifications au cours de l'histoire.

Cette histoire remonte à plus de deux cent ans avant l'émergence de la physique statistique. À la Renaissance, la probabilité était associée à l'*opinion* en opposition à la connaissance, vraie, absolue et certaine. Une opinion était « probable » si elle était approuvée par une autorité. Au Moyen-Âge et à la Renaissance, les preuves provenaient des témoignages des témoins et des textes ou paroles des autorités, bref des *gens*. L'expérience devait se conformer à ces témoignages et non l'inverse. Ce qui manquait était des preuves des *choses*. Une preuve des choses est différente d'une impression ou de données empiriques. Tandis qu'aujourd'hui l'on conçoit les livres et les gens, qu'ils soient des autorités ou des témoins, comme tenant lieu de preuves plus fondamentales, en ce sens où l'on est en droit de s'attendre qu'ils se réfèrent et s'appuient sur des faits, autrement dit sur des choses, à la Renaissance on procédait plutôt à l'inverse, les choses pouvant compter pour preuve en autant qu'elles s'accordent avec les témoignages et les livres. On peut situer l'origine de ce retournement dans le changement méthodologique face à l'expérience qui s'est opéré à la révolution scientifique. (Pour plus de détails, voir Hacking 1975.)

La tradition classique des probabilités tire son origine dans les années 1660 où émerge, donc, une nouvelle notion, soit celle « preuve interne », comme preuve provenant des choses, en distinction à une notion plus ancienne, celle de « preuve externe », comme les témoignages et les livres.² Dans cette tradition classique, issue principalement d'une rationalisation des « jeux de chance », la notion de probabilité est conçue comme une mesure du « degré de certitude de nos croyances » dans un domaine *fini* de possibles. À cette époque les auteurs croient au déterminisme absolu des phénomènes naturels et l'interprétation qu'ils favorisent, bien que pouvant être qualifiée de « subjective », en ceci qu'elle réfère à nos connaissances partielles, ne va pas aussi loin dans cette mouvance que les auteurs contemporains

² Cette distinction est très semblable à celle de Russell (1948) entre connaissance « par expérience » (« *by acquaintance* ») et connaissance « par description » (« *by description* »).

pour qui les croyances peuvent être basées sur des opinions voire des caprices. L'impact des probabilités classiques, comme branche des mathématiques appliquées, sur les sciences et la philosophie demeure somme toute négligeable. La seule règle alors disponible afin d'attribuer des valeurs spécifiques de probabilité consistait en le principe de raison insuffisante ou le principe d'indifférence, comme quoi tous les cas possibles se voient attribués la même probabilité, une situation d'équiprobabilité dirions-nous aujourd'hui. Mais la régularité dans les fréquences des phénomènes de masse³ a mené à cette idée alternative, à la base de l'interprétation fréquentiste, que les probabilités étaient sans doute l'expression d'une régularité naturelle et non une simple représentation de nos croyances.

Ce serait cette reconnaissance d'une régularité dans les choses à partir de laquelle des inférences inductives peuvent être tirées qui aurait rendu « possible », avec Hume, le problème sceptique de l'induction (Hacking 1975). Chez Hume (1739 : 174-204), le concept de probabilité en est un protéiforme (Hacking 1978 ; Landemore 2004). Il fait partie des trois types de rationalité, ou mieux, d'une distinction entre trois types de preuves : celui de la connaissance issue de démonstrations, celui des preuves issues de la relation de cause à effet, et celui provenant de probabilités. Dans le cas où une cause semble montrer plusieurs effets différents, voire « contraires », alors il peut s'agir d'une « probabilité de causes », où il y a contingence de causes, la « vraie » cause étant inconnue, où des causes contraires peuvent intervenir, ou encore il peut s'agir d'une « probabilité des chances » où ces différents effets présentent autant de chances égales, et sont donc équiprobables. Ces chances « signalent une indifférence subjective de l'esprit plutôt que l'existence d'un hasard objectif » (Landemore 2004 : 25). Cette indifférence peut découler d'un manque d'information quant à la « vraie » cause, information qui dissoudrait la variété des effets par une variété de causes. Mais ce qui devient vraiment différent, novateur avec Hume, est la possibilité, autrefois inconcevable, de passer de la probabilité à la preuve, c'est-à-dire qu'une croyance « probable » pouvait, par l'accumulation de preuves, devenir objet de connaissance⁴.

Jusqu'où cette approche des probabilités et de la connaissance générale a pu influencer, voire rendre possible, ce qu'on pourrait appeler le traitement probabiliste des phénomènes naturels demeure toutefois obscur. J'ai déjà signalé (Chapitre 3) que l'idée d'utiliser des arguments statistiques étaient « dans l'air » au milieu du XIX^e siècle en ce qu'elle fut mobilisée par plus d'un scientifique, de manière relativement rapide et avec une approbation assez bien répartie. Les exemples les plus fameux sont sans

³ Où un grand nombre d'agents semblables sont impliqués.

⁴ Le « scepticisme mitigé » de Pierre Gassendi (1592-1654), en opposition avec l'épistémologie de Descartes, est aussi un exemple d'une approche insistant sur la probabilité plutôt que sur la certitude (Osler 2000 : 165).

aucun doute le darwinisme et la théorie cinétique de la chaleur, qui allait devenir la mécanique statistique, Boltzmann ayant d'ailleurs manifesté beaucoup de sympathie pour la théorie de l'évolution par sélection naturelle (Prigogine & Stengers 1979 ; von Plato 1994 ; Cercigani 1998). Les premières applications du calcul des probabilités aux traitements des phénomènes thermiques remontent à August Krönig (1822-1879) en 1856 et à Clausius en 1857, pour l'essentiel avec le calcul de valeurs moyennes. Si, à partir de cette période, la physique semble faire un emploi « décomplexé » des probabilités, il n'en reste pas moins que la justification de cet emploi ainsi que leur interprétation fait l'objet d'une grande ambiguïté. En calculant la distribution des vitesses des molécules dans un gaz, Maxwell (1860) n'a pas tranché la question de savoir s'il s'agissait là d'une approximation, d'une moyenne ou d'une expectative. Pour sa part, Boltzmann (1868) interprète la distribution de Maxwell autant comme le temps relatif durant lequel une molécule (idéalisée) possède une certaine vitesse que comme le nombre relatif de molécules ayant cette vitesse. Cette association, tout de même problématique, est à la base de l'hypothèse ergodique, dont l'introduction est motivée par la possibilité de calculer les valeurs moyennes de quantités physiques de systèmes à plusieurs degrés de liberté, comme un gaz.

Ces difficultés traversent aussi, jusqu'à un certain point, les mathématiques car les similitudes conceptuelles avec la physique sont frappantes. Cette dernière offre en effet un « habitat naturel » (von Plato 1994 : 11) à la théorie des probabilités, notamment avec ses concepts d'espaces des états mécaniques, de temps continu et de processus markovien, où l'état du système n'est déterminé que par son prédécesseur immédiat et non par son histoire, ce qui fait écho au déterminisme mécanique (où l'état d'un système, à un certain instant, peut être déterminé uniquement à partir de son état à un instant antérieur arbitrairement proche et des lois mécaniques). En 1920, von Mises a tenté d'offrir une théorie ergodique détachée de la mécanique classique où le concept de probabilité y est défini avec la notion de hasard, comme limite d'une fréquence relative d'une séquence aléatoire. Ses travaux ont été poursuivis par von Neumann, qui a développé le formalisme de la mécanique quantique avec le concept d'espace hilbertien. Ce formalisme permet de caractériser un système physique par un vecteur, le vecteur d'état, dans cet espace mathématique où chacun des axes correspond à un état possible, en sorte que la projection de ce vecteur sur l'un des axes définit une probabilité de mesure pour cet état. Ainsi, les probabilités ont été incorporées très tôt dans le formalisme de la théorie. La preuve de von Neumann quant à l'impossibilité principielle d'inclure dans ce formalisme des « variables cachées », ce qui permettrait en quelque sorte d'identifier de « vraies » causes, a poussé la physique à franchir un pas décisif vers l'indéterminisme (pour plus de détails, voir Jammer 1974 ; von Plato 1994 ; Lindley 2007).

Les développements de la physique ont donc eu une influence profonde sur la conceptualisation des probabilités. Les travaux en lien avec l'hypothèse ergodique en physique statistique ont considérablement changé la théorie proprement mathématique des probabilités. La tradition classique

jouait surtout sur le terrain des applications associées à des cas simples, des domaines finis de cas équiprobables. Mais l'expérience ne montre pas ce genre de symétrie absolue et un changement vers une conception moderne des probabilités s'est opéré au tournant du XX^e siècle avec l'introduction d'événements *infinis*. Émile Borel (1871-1956), influencé par les traitements de problèmes physiques par l'astronome Hugo Gylden (1841-1896) et par Poincaré, fut le premier à s'investir dans ce changement conceptuel. David Hilbert et Richard von Mises ont bien sûr apporté des contributions significatives à la théorie moderne des probabilités. Mais son axiomatisation n'est venue qu'avec Andreï N. Kolmogorov (1903-1987) en 1933. Elle se présente, brièvement, comme suit : un *espace de probabilité* est tripartite tel que $[\Omega, \Lambda, P]$, où Ω est un ensemble non vide et fini, soit « l'ensemble universel » définissant l'ensemble des possibles, Λ est l'algèbre booléenne de tous les sous-ensembles de Ω , et P une fonction de Λ , appelée « fonction de probabilité ». Deux événements, A et B, mutuellement exclusifs, répondent à l'axiomatique suivante : (i) la *non-négativité*, définissant l'impossibilité d'une probabilité négative, de sorte que $P(A) \geq 0$; (ii) la *normalisation*, où il va de soi que la probabilité d'occurrence de l'un ou l'autre des possibles est égale à l'unité, tel que $P(\Omega) = 1$; et (iii) l'*additivité finie*, qui veut que la probabilité d'occurrence de A ou B est la *somme* des probabilités d'occurrence de A et de B⁵.

Influencé par les développements de la physique, les lois statistiques et probabilistes, « de pis-aller destinées à combler les insuffisances de nos moyens de connaissance, [sont ainsi] devenues des lois physiques autonomes » (Barberousse 2002 : 80). Mais cette influence n'est pas à sens unique. Car la théorie moderne des probabilités a su s'affranchir du domaine de l'expérience en faisant appel à la notion d'infini, alors que son développement a aussi contribué au formalisme d'une de ces théories fondamentales, la mécanique quantique. Sans doute un signe de la forte intrication entre ces deux disciplines (voir Gingras 2001 ; Dorato 2005).

5.1.2 Interprétations

Comme dans la plupart des discussions portant sur des concepts scientifiques, en simplifiant un peu voire beaucoup, une approche ontologique, généralement considérée comme « objective », ainsi qu'une approche épistémologique, plus près d'une conception « subjective », peuvent être distinguées. L'idée serait que le monde est indépendant de l'esprit humain mais que les discours sur le monde ne le sont pas. Par exemple, le micro-état, défini comme l'ensemble des variables mécaniques pour toutes les molécules, en mécanique statistique, n'est pas en principe inaccessible, mais le traitement statistique et

⁵ Formellement : $P(A \cup B) = P(A) + P(B)$ pour tout A, B $\in \Lambda$ tel que $A \cap B = \emptyset$.

probabiliste qui lui est réservé découle plutôt, en partie du moins, de nos limitations cognitives et expérimentales. Aussi, l'état d'un système quantique est représenté par une fonction censée contenir toute l'information pour le caractériser et ne permet que des prédictions sur un ensemble de résultats de mesure possibles auxquels on assigne des probabilités. (Bien entendu, ces descriptions font l'objet de débat.) Bien qu'essentielle, cette distinction n'épuise pas toute la richesse et la complexité du concept de probabilité.

À la question qu'est-ce que l'interprétation d'un concept ou d'une théorie doit précéder celle de savoir qu'est-ce qu'un concept, ou une théorie. Un dictionnaire donne comme définition d'un concept quelque chose comme une « représentation mentale générale et abstraite d'un objet ». Mais il y a d'autres propositions à ce qu'*est* un concept, comme celle les identifiant à des capacités d'agents cognitifs ou celle qui insiste sur l'aspect abstrait de la précédente définition en les présentant comme les constituants des propositions (voir Margolis & Laurence 2014). C'est pourquoi les concepts sont généralement considérés comme les constituants, les unités de la pensée. En ce sens, un concept n'est ni vrai ni faux. Selon l'intuition frégréenne, les concepts peuvent être compris sur le modèle des expressions fonctionnelles du langage mathématique, telle que « $x + 2 = 1$ » qui peut être vraie ou fausse si l'on assigne une valeur à x . Lorsqu'un concept est inséré comme argument au sein d'une expression formée d'un sujet, alors cette expression peut être considérée complète et susceptible de recevoir une valeur de vérité (Laurier 1980 ; Panaccio 2011). Un concept peut donc être le prédicat d'un sujet, mais il peut aussi recevoir une désignation par un symbole ou un mot, et une dénotation ou un référent, en ce qu'il désigne ou renvoie à quelque chose d'autre, réel ou non⁶. Ainsi, la signification ou le sens attribué à un concept est donné par son *intension*, soit l'ensemble de ses propriétés et relations, tandis que l'ensemble des choses, réelles ou non, pouvant y être subsumées constitue son *extension*, son domaine d'applicabilité (par ex., le concept de fantôme a une intension, un sens, et une désignation par un mot, mais pas d'extension).

Il va donc de soi que certains concepts sont complexes en ce qu'ils sont composés de concepts plus primitifs ou plus simples. Mais en jouant ce rôle d'unités constitutives de la pensée, les concepts ne peuvent se résumer à des termes linguistiques, certains concepts tirent des mots leur portée représentationnelle, les « images » que l'on peut s'en faire, et d'autres non. La détermination de la signification d'un concept peut donc être particulièrement difficile. À cela il faut ajouter que cette signification peut être affectée par le contexte dans le lequel le concept est utilisé, que ce soit au sein

⁶ Il me paraît utile de définir un référent comme une « dénotation réelle » ; par exemple, le concept « nombre premier » a plusieurs dénotations (1, 3, 5, etc.), constituant son extension, mais elles ne sont pas réelles, au sens empirique du terme, donc il n'a pas de référent physique. Bunge (1967a) soutient plutôt que la dénotation est l'union de la désignation et de la référence.

d'un énoncé, d'une théorie ou même d'un modèle. Mais si la signification des concepts est susceptible de changer, est-ce que la valeur de vérité des énoncés, théories et modèles change aussi ? Sous quelles conditions ces deniers peuvent-ils être considérés comme vrais ? C'est ici qu'entre en jeu l'interprétation.

Selon Russell (1948), la procédure d'interprétation consiste à établir une signification stable à une classe de termes indéfinis. Cette signification peut être donnée par une définition nominale ou ostentatoire, comme quoi l'interprétation peut être de type logique ou empirique. C'est précisément par ce procédé que les termes théoriques sont susceptibles de recevoir une valeur de vérité. Ainsi, une interprétation empirique offre un lien, une « relation de similarité » (Teller 1995 : 5) censée tenir entre un modèle ou une théorie et la réalité qu'ils tentent de décrire. Pour reprendre les termes d'Omnès (1994 : 237), elle tente de réconcilier l'« alpha du fait » avec l'« oméga de la théorie ». Ce qui ne veut pas dire qu'une nouvelle interprétation mène à de nouvelles prédictions, mais plutôt elle se tient sur le plan de l'inobservable, se limite à des « variables cachées empiriquement superflues » (van Fraassen 1991 : 243). À titre d'exemples, l'interprétation de certains phénomènes électriques à partir de la notion de « courant électrique » ne sera pas affectée par les résultats expérimentaux obtenus aux moyens d'une pile et d'un ampèremètre ; ou encore, l'interprétation de l'entropie comme une « mesure du désordre » ne change pas les prédictions de la mécanique statistique à propos des phénomènes thermiques⁷.

En tant que partie d'un système théorique les termes indéfinis doivent satisfaire à certaines contraintes logiques ou mathématiques, ce qui ne se limite donc pas à ce qui est empirique et vérifiable, mais leur interprétation ne doivent pas invalider (rendre fausses) les relations logiques ou mathématiques au sein desquelles résident ces termes. On s'attend ainsi à ce qu'une interprétation du concept de probabilité n'invalide pas le calcul des probabilités. Mais si deux interprétations satisfont ces contraintes, alors, selon Russell (1948 : 256), le choix revient à une question « de goût et de commodité ». On peut emprunter le raisonnement de Russell en affirmant que les discussions quant à la « bonne » interprétation d'un terme théorique, comme avec la probabilité ou l'entropie statistique, s'inscrivent dans un « système » plus large que le système théorique au départ, comme un « système » métaphysique ou épistémologique, par exemple. Mais on peut aussi avancer qu'une « bonne » interprétation favorise la compréhension d'un concept généralement abstrait en le ramenant à des concepts plus intuitifs. C'est pourquoi le terme « interprétation » est aussi défini comme « action

⁷ Plus spécifiquement, il y a plusieurs fonctions de distributions (ou d'opérateurs densité) compatibles avec des paramètres macroscopiques donnés, de sorte qu'il faut se donner un critère afin d'effectuer un choix. Ce critère est appelé le *postulat de maximisation de l'entropie statistique*, qui stipule que l'on doit choisir la fonction de distribution maximisant l'entropie statistique. Mais ce choix demeure le même si l'on décide de l'interpréter comme le choix du macro-état le plus « désordonné » ou le plus « probable » ou le plus « dispersé ».

d'expliquer ». C'est ce que tente de faire, par exemple, avec plus ou moins de succès, la proposition visant à interpréter le concept d'entropie par celui de « désordre » ou celle des probabilités par les « propensions » (discuté plus loin). Une « bonne » interprétation devrait aussi, minimalement, être précise, non-ambiguë, non-circulaire et faire appel à des concepts primitifs qui soient clairs et bien compris.

Il y a cinq interprétations principales des probabilités selon Hájek (2012), soit les interprétations classique, logique, fréquentiste, propensionniste et subjectiviste (voir aussi Carnap 1950 ; Galavotti 2008). Celui-ci présente d'abord les critères que devrait rencontrer une interprétation « adéquate » selon Salmon (1966) : l'*admissibilité*, comme quoi l'interprétation d'un système formel rend les énoncés qu'il contient vrais ; la *vérifiabilité*, identifiant une méthode d'attribution et de vérification de valeurs de probabilités, et enfin l'*applicabilité*, comme quoi les probabilités peuvent être utiles, et peut-être même seraient-elles le « guide la vie » selon l'expression consacrée. Il y a en effet, tel que mentionné, une multitude de situations où sont utilisés les probabilités et les raisonnements qui s'y appuient ne sont pas qualifiés d'irrationnels, au contraire. La question est de savoir à quel genre de raisonnement et d'inférence le concept de probabilité devrait s'appliquer. Hájek identifie ensuite, donc, une liste de critères d'applicabilité qui n'est ni exhaustive ni nécessaire, mais il serait tout de même surprenant qu'une interprétation particulière se trouve en porte-à-faux de chacun d'entre eux : la *non-trivialité* (par ex. où seules les valeurs 0 et 1 sont permises), l'*applicabilité aux fréquences* (relation entre probabilités et fréquences relatives), l'*applicabilité à la croyance rationnelle* (degré de croyance d'agents rationnels agissant selon les probabilités, par ex. en étant confiant en l'occurrence d'un événement très probable), l'*applicabilité aux inférences ampliatives* (un atout serait qu'elle puisse identifier les bonnes ou mauvaises inférences par rapport aux inférences strictement déductives), et l'*applicabilité à la science* (par ex. en mécanique quantique et en mécanique statistique). Il n'est pas évident que ces critères soient minimalement compatibles, ni qu'une interprétation puisse tous les satisfaire. Il demeure aussi possible que des desiderata métaphysiques soient souhaitables, comme celui liant probabilité et modalité (Chapitre 2).

Je vais me référer à cette liste de critères lorsque je présenterai quelques pistes de solutions à l'interprétation des probabilités, en affirmant non pas la dualité des probabilités *per se* mais plutôt celle des modes d'attribution empirique (critère de vérifiabilité) de valeurs ou de mesures de probabilités (section 2.3). Je propose alors de considérer les critères d'applicabilité aux fréquences et aux croyances rationnelles comme étant nécessaires à l'attribution de valeurs de probabilité, et en ce sens le concept de probabilité ne peut se réduire à l'un de ses aspects, « objectif » et « subjectif ».

L'interprétation *classique* assigne une valeur de probabilité en l'absence de preuve quant à une occurrence particulière (par ex. les conditions initiales sont inconnues malgré des lois déterministes) ou dans le cas de preuves symétriquement réparties (par ex. un dé). Dans de telles circonstances, la probabilité est partagée également parmi tous les résultats possibles, de sorte que la probabilité classique d'un événement est simplement la fraction du nombre total de possibilités de l'événement l'axiome 2 est alors trivial. Plusieurs objections peuvent être faites contre cette interprétation. D'abord, quel est le critère permettant de déterminer que des événements sont « du même type », qu'ils sont également probables ? Cela renvoie au problème de la classe de référence (discuté plus loin). Ensuite, la notion de « également probable » semble circulaire puisqu'elle peut signifier « également possible ». Aussi, il est loin d'être clair qu'une possibilité puisse prendre différent degré, comme le manifeste la quantification du calcul des probabilités. Enfin, cette interprétation ne permet pas des probabilités dans l'ensemble des réels, telles que $1/\sqrt{2}$, ni des ensembles infinis, comme l'exige une théorie scientifique comme la mécanique quantique. (Pour plus de détails, voir Galavotti 2008 ; Håjek 2012.)

L'interprétation *logique*, dont l'étude la plus systématique est celle de R. Carnap (1950 et 1966), reprend l'idée de l'interprétation classique voulant que les probabilités puissent être déterminées *a priori* par un examen du spectre des possibilités. Mais elle effectue une généralisation de deux importantes façons : les probabilités peuvent recevoir des poids inégaux, et celles-ci peuvent être calculées peu importe les preuves. Un élément important de cette interprétation est qu'elle fournit un cadre à l'induction, car elle englobe le degré de confirmation c qu'une preuve e confère à une hypothèse h , tel que $c(h, e)$, et fournit ainsi une théorie générale de l'inférence. De sorte que, la probabilité étant une implication partielle, la déduction est un cas extrême d'une implication de probabilité égale à l'unité. Toutefois, si la fonction $c(h, e)$ permet une confirmation systématique d'une preuve et d'une hypothèse, pour autant que celles-ci soient valides, cette fonction même n'a pas de justification hormis celle de la simplicité, car il existe plusieurs autres fonctions de confirmation.

L'interprétation *fréquentiste* a deux versions, la version opérationnaliste ou finitiste (séquence finie) et la version hypothétique (fréquence limite d'une séquence infinie). La première définit la probabilité d'un attribut A d'une classe de référence finie B comme la fréquence relative d'occurrences réelles ou observées de A dans B. Elle présente donc certaines similarités avec l'interprétation classique, en ce sens qu'elle définit des membres équiprobables d'un ensemble, mais elle s'en détache aussi puisqu'elle considère les résultats « réels » (c'est-à-dire observés) et non pas seulement « possibles ». Elle est donc sympathique aux scrupules empiristes. Cependant, comme elle donne une définition opérationnelle des probabilités, elle se heurte au *problème du cas unique* : par exemple, une pièce de monnaie qui n'a jamais été tirée n'a pas de probabilité de tomber sur l'une ou l'autre de ses faces, ou encore pire, si elle n'est tirée qu'une seule fois, elle a nécessairement une probabilité de 1 ou 0. La seconde version fait appel

aux contrefactuels, à une fréquence relative limite, en imaginant quelle *serait* la fréquence relative *si* la séquence était étendue jusqu'à être infinie. L'empirisme semble ici être laissé de côté.

Or, cette fréquence relative limite est toujours relative à une classe de référence, qui demeure difficilement définissable et identifiable, d'où le *problème de la classe de référence* (Kyburg 1975 ; Hájek 2012 ; Handfield 2011). Car pour définir cette classe, il faut qu'un ensemble de cas « suffisamment similaires » ou « suffisamment pertinents dans un contexte donné » soit identifié avec « suffisamment de précision ». Or il ne semble pas y avoir de principe clair ou suffisamment consensuel ou formel déterminant ces paramètres et donc la classe de référence. Et même si tel était le cas, la séquence à l'intérieur de la classe de référence peut facilement être sous-déterminée si la correspondance entre le *definiens* et les *definienda* de la classe de référence n'est pas clairement établie. Par exemple, la probabilité attribuée à la défaillance d'un réacteur nucléaire peut être évaluée avec la fréquence relative des réacteurs défectueux par rapport à tous les réacteurs nucléaires ; mais la classe « tous les réacteurs nucléaires » est plutôt ambiguë : sont-ce ceux du même modèle, ou installés dans le même genre de centrale, depuis un certain temps, etc.? (Pour plus de détails, voir von Plato 1994 ; Hájek 1996 ; Handfield 2011.)

Selon l'interprétation *propensionniste*, sans doute anticipée par Charles Sanders Pierce, la probabilité est une propension physique, une disposition, ou une tendance pour un phénomène physique donné à donner un certain résultat ou une certaine fréquence relative. Elle tient donc de la doctrine ontologique du réalisme générique comme quoi il y a une réalité qui se donne au sujet et indépendante de celui-ci. Cette approche est surtout motivée par le désir de résoudre le problème du cas unique rencontré principalement dans les phénomènes quantiques. Cependant, une description des probabilités est insuffisante à *prouver* l'existence de telles propensions, qui demeurent *per se* inobservables. Stipuler ce qu'est un fantôme est insuffisant à prouver que les fantômes existent. En effet, si la description typiquement probabiliste de certains phénomènes semble avoir recours à des termes comme les fréquences (par ex. le tir d'un dé), ce n'est pas le cas des propensions. En revanche, ce concept offre une explication de la stabilité de certaines de ces fréquences. Une autre critique consiste à soutenir que la position voulant qu'un arrangement expérimental ait une tendance à donner une certaine fréquence relative, en raison précisément d'une propension, présuppose une certaine uniformité de la nature. Car autrement la raison première de la stabilisation de la fréquence relative ne pourrait être attribuée à la propension, mais à des facteurs externes comme le contexte expérimental. Cette critique, que l'on peut ramener à Hume, semble viser l'applicabilité aux sciences, mais elle alors est moins effective si l'on pose que l'uniformité de la nature est une condition de possibilité des sciences (Einstein 1949). Enfin, l'applicabilité aux sciences est mise à mal par la non-testabilité des énoncés propensionnistes pour un cas unique, ce qui leur conférerait un aspect plus métaphysique que scientifique, bien qu'il ne faille pas

confondre cette condition d'applicabilité et une certaine exigence de calquer l'interprétation sur les conditions de validité des sciences. (Pour plus de détails, voir Popper 1959 ; Settle 1975 ; Parker 2001 ; Belnap 2007.)

L'interprétation *subjectiviste* se laisse résumer par le slogan « la probabilité est le degré de croyance »⁸. Bien entendu, plusieurs interprétations sont alors disponibles selon l'agent considéré et le problème revient donc à le définir. Le *subjectivisme non contraint* n'impose aucune contrainte à l'agent, tandis que le *subjectivisme rationaliste* impose un agent « rationnel ». Ces thèses se déclinent en plusieurs versions, dont le *bayésianisme*, selon lequel les croyances rationnelles doivent obéir aux axiomes de la théorie des probabilités et qu'elles doivent être mises à jour par « conditionnalisation » à la lumière de nouvelles informations. Cette mise à jour s'appuie sur le théorème de Bayes décrivant, à partir de la probabilité conditionnelle, la probabilité P d'un événement A sachant B telle que

$$P(A|B) = [P(B|A) P(A)] / P(B).$$

La question revient alors à définir ce qui constitue des raisons suffisantes pour déterminer les probabilités aprioriques, $P(A)$ et $P(B)$. Il ne faudrait toutefois pas confondre trop rapidement *bayésianisme* et *subjectivisme* (Carnap 1950 ; Bitbol 1997 ; Bricmont 2001). D'une part, il n'y a pas de connexion logique entre ces deux thèses, car l'on peut croire qu'il y a des faits objectifs et soutenir que les probabilités constituent un « outil utile » lorsque nos connaissances sont incomplètes. Une évaluation rationnelle des probabilités relativement à un contenu d'information préalable est une opération tout ce qui a de plus objectif. D'autre part, il faudrait voir à ne pas confondre le monde et ce qu'on en connaît : des connaissances incomplètes n'implique par une ontologie déterministe ou indéterministe (par ex. les jeux de hasard sont compatibles avec le déterminisme). (Pour plus de détails, voir Norton 2007 ; Galavotti 2008 ; Elga 2010 ; Handfield 2011.)

On constate donc que la théorie des probabilités, malgré son utilité et la confiance qu'on lui porte, mais surtout malgré son axiomatisation rigoureuse, est soumise à une pléthore d'interprétations faisant intervenir un ensemble de questions et de problèmes épistémologiques avec lesquels les théories scientifiques doivent aussi conjuguer. Dans la discussion qui suit, je soutiens que toutes ces interprétations, qui insistent sur l'un ou l'autre des aspects de la méthode d'attribution de valeurs de probabilité, ne voient qu'un aspect du problème : en tentant de réduire le schème dualiste de la connaissance, le clivage objectif-sujetif, à l'un de ses termes, c'est précisément cette méthode

⁸ Elle semble donc tenir de la doctrine épistémologique de l'internalisme, comme quoi la connaissance ne peut reposer sur des hypothèses non justifiées et cette justification est nécessairement interne et subjective.

d'attribution de valeurs de probabilité qui devient inopérante et par conséquent le critère d'applicabilité n'est pas satisfait.

5.1.3 Discussion

Selon von Mises (1928), les probabilités tirent leurs origines d'où on les introduit : « *probability in, probability out* ». Qu'est-ce que cela signifie ? En termes simples, que les probabilités rentrent par où on les introduit et qu'elles sortent si on les y conduit. Plus précisément, cela signifie que l'introduction de probabilités dans un raisonnement et à plus fortes raisons dans une théorie est *stipulée*, puisque les probabilités ne peuvent être ni observées ni être déduites d'énoncés exempts de probabilités. Cela signifie aussi que l'élimination du caractère probabiliste d'un raisonnement est logiquement impossible une fois que l'introduction des probabilités est opérée (à moins de trivialement considérer une probabilité nulle ou égale à l'unité comme étant simplement une absence de probabilité). En somme, une théorie mathématiques des probabilités est fermée sur elle-même : elle ne peut appuyer ses estimations probabilistes que par une autre estimation probabiliste. Il faut donc distinguer clairement le *calcul* des probabilités de l'*attribution* de valeurs de probabilités.

Peut-on *prouver* cette affirmation quant à la non-observabilité et à la non-déductibilité des probabilités ? La question est similaire, selon moi, à celle quant à la possibilité de déduire l'existence d'une entité à partir de prémisses ne stipulant pas cette existence, à celle aussi quant à la possibilité d'obtenir une conclusion normative de prémisses descriptives, ou encore à la possibilité de déduire un comportement indéterministe de certains systèmes à partir de prémisses spécifiant un comportement déterministe. La différence en est une de pertinence et de compétence. La question est ainsi de savoir s'il y a un véritable fossé conceptuel et donc si les précédentes distinctions sont irréductibles à l'un des termes. L'argumentation qui suit servira à appuyer, à tout le moins, cette affirmation établissant la précédente distinction, dans l'objectif d'un point de vue original sur les probabilités, d'en jeter les bases pour une nouvelle interprétation dans le cas de probabilités appliquées au comportement de systèmes naturels, à des événements dits réels, en ce sens où leur description est susceptible d'être corroborée ou falsifiée.

On peut avoir de bonnes raisons d'attribuer une probabilité à un événement ou encore une mesure de probabilité à un ensemble d'événements, mais l'attribution ne fait tout simplement *pas* partie du calcul des probabilités. Il n'y a aucune règle dans la théorie des probabilités statuant sur la procédure à suivre pour l'identification d'événements probabilistes *dans les faits*. C'est-à-dire qu'il n'y a rien pour encadrer précisément l'attribution des probabilités, hormis bien sûr un cadre général défini par les

axiomes comme ceux de la non-négativité et de la normalisation. Ce cadre général autorise une infinité de mesure de probabilités à une infinité d'événements. Ces règles d'attribution sont à trouver plutôt dans l'interprétation des probabilités où il est possible d'établir une procédure opératoire. Par exemple, la stabilisation d'une fréquence relative à la suite d'un ensemble d'événements similaires peut être une bonne raison d'attribuer une probabilité à un événement particulier. Mais il ne va pas de soi que cette fréquence relative puisse être raisonnablement *identifiée* à la probabilité de l'événement, ni qu'il y ait des critères justifiant l'inférence inductive de l'un à l'autre. Pour paraphraser Caves, Fuchs & Schack (2007 : 260), les fréquences observées lors d'essais répétés ne sont pas des probabilités, ce sont des *faits*. L'identification pourrait être justifiée, encore que cela demeure tout à fait discutable, dans le cas d'une séquence infinie, mais il est évident que ce cas ne peut être instancié. Il ne va pas de soi non plus que des quantités normalisées sans dimension (par ex. le ratio de l'aire du centre d'une cible par celle de la cible totale) puissent aussi être identifiées à des probabilités, et qu'au surplus elles représentent des quantités *physiques*. La théorie des probabilités est compatible avec une pléthore d'interprétations même si des arguments philosophiques favorisent certaines interprétations plutôt que d'autres. Donc, la procédure mathématique et déductive permettant d'inférer des valeurs de probabilités ne détermine pas la procédure empirique permettant d'identifier des événements et de leur attribuer une probabilité.

Ainsi le calcul des probabilités se place sur le plan formel et logique, c'est-à-dire non factuel, en sorte qu'en l'absence d'interprétation il n'offre aucune signification des termes en dehors des relations logiques et mathématiques qu'il établit. En revanche, l'attribution d'une valeur de probabilité aux phénomènes naturels procède de justifications qui ne se limitent pas aux règles logiques et mathématiques : elles dérivent d'« inférences informelles » (Bunge 1967a : 90). La situation, ici pour les probabilités, est similaire, nonobstant la dimension éthique, à celle rencontrée dans la « monétarisation » de la nature, c'est-à-dire la procédure visant à donner une valeur monétaire à un écosystème ou à l'une de ses parties (Maris 2006 ; Maresca 2008) : le jugement pondérant les valeurs *déjà établies* est basée sur des règles arithmétiques simples et claires, mais il en va autrement de la justification et de l'attribution de cette valeur monétaire, en amont de ce jugement, pour un système naturel donné. Qu'un certain concept ait telle signification et qu'il suive logiquement d'un second concept est une chose, mais que l'un ou l'autre de ces concepts possède une extension pouvant recevoir un référent empirique, observable ou susceptible de revoir une définition ostentatoire, en est une autre. En d'autres mots, comme le fait remarquer à juste titre Bitbol (1997 : 98), il faut voir à ne pas confondre la méthode et la justification des prémisses qu'on utilise pour la mettre en œuvre.

5.1.3.1 SCHÈME OBJECTIF-SUBJECTIF

Rendre compte des justifications appuyant ce genre d'attribution de probabilité peut s'avérer excessivement ardu et fait intervenir tout un réseau de problèmes philosophiques particulièrement importants pour la philosophie des sciences. D'abord, l'attribution d'une valeur de probabilité qui n'est pas simplement stipulée n'a sans doute pas le même caractère d'évidence ou d'immédiateté que peut avoir, par exemple, l'attribution de la couleur bleue à un objet particulier. Ensuite, les justifications de ces attributions doivent faire face aux mêmes problèmes que ceux reliés aux conditions d'une corroboration et d'une falsification d'une théorie, c'est-à-dire, de manière très générale, qu'il faut établir les conditions permettant d'identifier un fait ou un ensemble de faits comme faisant partie ou non d'une classe nécessairement définie par des énoncés généraux, voire universels. À titre d'exemple, l'inférence passant de l'observation d'une fréquence relative à une probabilité présente tous les éléments du problème de l'induction (surtout pour le fréquentisme hypothétique). Un problème évidemment relié à celui de la classe de référence (vu précédemment). Enfin, il est évident que ces justifications dépendent de la définition et de l'interprétation que l'on peut donner au terme « probabilité ». Au premier chef, la distinction entre deux concepts de probabilité, discutée par Carnap et Hacking et recoupant le couple objectif-subjectif, joue un rôle majeur. Typiquement, pour certains il est justifié de s'appuyer sur l'observation pour procéder à une telle attribution, ce qui semble être objectif (ou non), tandis que pour d'autres l'observation serait insuffisante ou injustifiée, voire impertinente. En insistant sur ce qui constitue une justification à ces attributions de probabilité, les problèmes épistémologiques associés aux probabilités et à leurs interprétations ressortent avec plus d'acuité.

En ce qui concerne le problème de l'induction appliqué aux fréquences relatives, par exemple, quelle justification permet d'assurer que le nombre d'occurrences soit suffisant pour supporter l'inférence à la probabilité à partir de cette fréquence ou que la marge d'erreur soit suffisamment petite ? Et est-il possible d'éliminer toute trace d'arbitraire dans ces opérations de fractionnement (la probabilité de *quoi* ?) permettant de dénombrer les unités individuelles et d'identifier les classes de références ? Je soutiens ici qu'il y a deux aspects essentiels à la méthode d'attribution empirique de valeurs de probabilités : un aspect objectif, associé d'une part à la mesure d'un système permettant d'établir une gamme de possibles et d'autre part à la caractérisation nomique des systèmes ; et un aspect subjectif (au sens faible, plutôt : arbitraire), associé d'une part à l'identification d'événements « suffisamment semblables » et d'autre part aux choix d'une mesure de probabilité (la plus commune étant la mesure uniforme menant à l'équiprobabilités des possibles). Cette interprétation implique l'irréductibilité de ces deux aspects en ce qui concerne l'application du calcul des probabilités et ne

suffit donc pas à trancher le schème dualiste de la théorie de la connaissance, celui de la dichotomie objectif-sубjectif.

Je soutiens donc que cette distinction entre calcul et attribution de probabilités, entre procédure mathématique et procédure empirique, est au cœur du débat interprétatif des probabilités quant à leur caractère objectif ou subjectif. Il y a cependant d'autres critères de classification des interprétations des probabilités étant donné, tel que discuté, la portée de ces interprétations en philosophie des sciences⁹. Ainsi, d'aucuns peuvent insister sur le caractère partiellement indéterminé, arbitraire, et (donc ?) « subjectif » de l'attribution de valeur de probabilités aux événements, aux faits, tandis que d'autres peuvent insister plutôt sur le caractère davantage « objectif » de la procédure déductive tirée de la théorie des probabilités. Mais il y a aussi un autre clivage objectif-sубjectif au sein de ce qu'on pourrait appeler l'« appareil probabiliste », qui inclut la procédure mathématique et la procédure empirique discutées précédemment. Il s'agit, justement, de la position face à une mesure probabiliste et à ses conséquences théoriques. En général, et tel que mentionné, pour un appareil probabiliste donné, donc pour un ensemble d'événements auquel est attribuée un ensemble de probabilités, la mesure appliquée à une trame factuelle, soit l'observation de l'occurrence d'un événement particulier comme la mesure d'un paramètre établissant une valeur particulière, est considérée comme un événement « certain », c'est-à-dire que l'on considère que cet événement s'est réellement produit même si avant la mesure il n'avait qu'une certaine probabilité de se produire. En ce sens, l'axiome de la normalisation (ii, section 2.1) ne signifie rien d'autre qu'une occurrence fait nécessairement, *ex hypothesi*, partie de la gamme des possibles. Une valeur de probabilité pour l'ensemble des possibles inférieure à l'unité signifie que cette gamme est incomplète, ce qui implique l'impossibilité d'attribuer une valeur de probabilité à certains événements, tandis qu'une valeur supérieure à l'unité signifie qu'un événement, un résultat préalablement défini comme impossible pourrait se produire (ce qui ne veut pas dire qu'il s'est effectivement produit). Ce qui serait, dans les deux cas, contradictoire et donc inacceptable.

L'aspect objectif des probabilités se manifeste, dans un premier temps, dans la description nomique ou mathématique des systèmes. L'objectivité à l'œuvre ici est donc celle des lois scientifiques, comme la dynamique, et celle du calcul (mathématique) des probabilités. Par exemple, *si* une loi est considérée valide, alors la description du système s'y appliquant est considérée objective, comme l'accélération gravitationnelle d'un corps n'est pas une question de préférence, une fois que sont établis les paramètres de la loi de la gravitation et qu'ils sont appliqués correctement à un corps particulier¹⁰ ; ou *si* une valeur

⁹ Par exemple, Bitbol (1997 : 90) présente le clivage d'inspiration kantienne entre contenu empirique et conditions formelles.

¹⁰ D'après une lecture kantienne, il y a aussi objectivité au sens où le système est conditionné par ce qui s'approche d'une application de la catégorie de la causalité.

de probabilités est correctement attribuée au système, alors les inférences strictement basées sur le calcul des probabilités sont objectives, au même titre que la probabilité d'occurrence conjointe d'événements indépendants de probabilité $\frac{1}{2}$ est nécessairement $\frac{1}{4}$. Si un système obéit aux lois de la mécanique classique, on ne dira pas que son évolution dans le temps est subjective, qu'elle dépend des préférences ou des états mentaux de tel ou tel individu, car elle est déterministe¹¹, donc déterminée par des paramètres qui n'ont rien de subjectif¹². Bien entendu, l'application de ces lois et règles doit être faite dans les règles de l'art et la possibilité de l'erreur n'est pas éliminée. On peut bien arguer que telle loi ne s'applique pas à tel système, ou que les valeurs des paramètres n'ont pas été correctement mesurés, ou que telle valeur de probabilité ne tient pas, ou que les événements ne sont pas vraiment indépendants, etc. Mais ce n'est pas le caractère objectif des lois qui est remis spécifiquement en cause.

L'aspect objectif des probabilités se manifeste, dans un second temps, dans la mesure empirique, c'est-à-dire, dans ce qui permet de corroborer ou de falsifier un énoncé, d'établir une gamme de possibles ainsi qu'une valeur de probabilité spécifique selon cette gamme. La mesure est donc entendue au sens large et pas seulement dans son acception quantitative¹³. L'important ici est de reconnaître que la satisfaction de l'axiome de normalisation ainsi que l'attribution d'une valeur de probabilité peuvent être vérifiées dans les faits. C'est-à-dire qu'il est possible d'établir, mais aussi de corriger, une gamme de possibles pour un phénomène donné, ainsi qu'un énoncé probabiliste. Par exemple, un septième résultat peut mener à une révision de la gamme ordinaire de possibles pour un dé, ou encore une valeur de probabilité $\frac{1}{2}$ attribuée à l'occurrence « face » peut être modifiée à la lumière de résultats ne donnant que des « piles ». L'objectivation d'un résultat de mesure est donc essentielle car elle permet une distanciation de ce qui est directement appréhendé par les sens par la formation d'un ensemble d'événements jugés similaires. Mais elle permet aussi de juger du « degré de croyance ». La croyance porte en effet sur des possibilités généralement déterminées de manière empirique. De plus, une fréquence relative n'est pas définie par rapport à cette croyance, mais par rapport aux résultats de mesure, par rapport aux événements, *simpliciter*. Un jugement définissant cette croyance peut bien sûr, ensuite, être porté sur cette fréquence relative et ainsi attribuer une valeur de probabilité. Mais ce jugement doit être distingué des résultats de mesure, des événements. En conséquence, la probabilité ne peut être réduite à la fréquence relative. Cette distinction permet d'éviter les conséquences du

¹¹ Earman (1986 : Chapitre 3) soutient plutôt que l'espace-temps de la physique newtonienne n'est pas un « environnement amical » au déterminisme laplacien. L'argument s'appuie principalement sur l'absence de vitesse maximale et n'implique pas *ipso facto* la mort du déterminisme.

¹² Quoique plus compliquée, la situation est similaire en mécanique quantique puisque l'équation de Schrödinger est déterministe. La question est davantage de savoir à *quoi* s'applique cette évolution déterministe.

¹³ À moins d'avis contraire, il s'agira ici d'une mesure au sens empirique et non théorique du terme, comme la mesure de Lebesgue.

problème de la classe de référence quant à l'objectivité des probabilités : différentes classes de référence, établies selon différentes fréquences, impliqueraient différentes valeurs de probabilités pour un même événement (Hájek 1997 ; Handfield 2011).

Pourtant, si les probabilités ne se réduisent pas aux fréquences, ces dernières n'en constituent pas moins une base d'inférence importante pour l'attribution de valeurs de probabilités. Il faut donc distinguer l'observation d'une fréquence de l'inférence s'y appuyant pour conclure à une certaine valeur de probabilité. L'inférence peut bien entendu être invalide ou injustifiée, mais la fréquence n'en demeure pas moins objective – il s'agit d'un fait. La séquence d'occurrences de valeurs particulières, pour un système donné et un protocole donné (par ex. le tir d'un dé), est objective et ne dépend pas des attitudes ou des préférences d'un quelconque sujet. Mais l'expectative de ces occurrences est autre chose et il va de soi que cette séquence puisse mener à plusieurs croyances différentes. Affirmer ainsi qu'un arrangement expérimental donné est « susceptible de produire » certaines occurrences ne renvoie pas à des croyances. La relative robustesse de ces occurrences pour des arrangements expérimentaux donnés implique plutôt l'objectivation des résultats de mesure permettant de définir la fréquence relative selon ces occurrences. Cela n'implique pas, en revanche, le choix d'une explication particulière pour une certaine stabilité dans ces occurrences, comme les propensions par exemple. L'objectivation des résultats de mesure est donc essentielle car autrement aucune fréquence relative ne pourrait être établie. Cela constitue l'aspect objectif des probabilités et plus particulièrement de la méthode d'attribution de valeurs de probabilités dans les faits.

La manifestation de ce caractère objectif est aussi présente dans l'usage courant des probabilités. Il ne fait aucun sens, par exemple, de dire que l'événement qui vient de se produire *a* une chance sur deux de s'être produit, mais on peut bien dire que l'événement *avait* une chance sur deux de se produire. Si, en revanche, une mesure n'épuise pas les possibilités, alors on considère que les connaissances sont incomplètes, mais pas nécessairement que le résultat de mesure est subjectif. On remarque bien ici la polarité entre « attente » et « constat » issue de la dissymétrie entre le passé et le futur au sein de l'appareil probabiliste. Ce serait dans l'attente que le caractère subjectif ou partiellement arbitraire se manifesterait avec plus d'acuité. Ainsi, s'il semble rationnel d'attribuer une probabilité de $\frac{1}{2}$ au tir d'une pièce de monnaie donnant le résultat « pile », il semble que ce ne soit pas le cas si l'on attribue *aussi* une probabilité de $\frac{1}{2}$ au tir qui vient d'être effectué et qui a effectivement donné le résultat « pile ». Le constat est susceptible de corroborer la croyance d'un quelconque sujet mais aussi de la falsifier, et c'est principalement dans l'erreur que l'aspect objectif de la probabilité se manifeste. Comment en effet affirmer que la probabilité découle uniquement du sujet si une mesure peut le contredire ? Le constat peut aussi faire l'objet d'un accord intersubjectif au sein d'une communauté de scientifiques. Tel que mentionné, le constat se manifeste dans la mesure qui permet de

déterminer la gamme des possibles, directement avec le dénombrement des occurrences, ou non avec la détermination des paramètres établissant le calcul des états inobservables compatibles (discuté plus loin).

La raison de cette polarité est sans doute que ce qui est mesuré est passé et ce qui est passé est « fixé », ce ne peut faire l'objet d'indétermination ou d'incomplétude, à moins que ce soit la mesure elle-même, donc nos connaissances, qui soit indéterminée et incomplète (Schaffer 2007). On touche là un problème difficile mais néanmoins crucial. Car nos connaissances du passé ne sont pas nécessairement « parfaites » ou « complètes », et elles ne sont pas nécessairement meilleures, au demeurant, que celles concernant l'avenir (Sklar 1993 ; Price 1996 ; Barrett & Sober 1999). La question n'est donc pas tant que certaines connaissances portent sur le passé et d'autres sur l'avenir, mais plutôt que ce qui est mesuré, au sens empirique du terme, et qui est donc par définition passé, offre un accès épistémique privilégié, quasiment assuré. Cela signifie qu'une mesure répond aux conditions générales permettant de corroborer une hypothèse ou simplement d'appliquer un concept ou d'en établir l'extension, à tout le moins partiellement. Ainsi, les occurrences d'un certain attribut, définissant une fréquence relative, finie et mesurée, au sein d'une certaine gamme de possibles ou classe de référence, n'ont pas de probabilité *au-delà* de celle définie par cette gamme, au risque de contredire l'axiome de la normalisation¹⁴. Si l'attribut est « pile » et la gamme {« pile », « face »}, l'occurrence mesurée, ou si l'on veut, observée, n'a pas de probabilité définie sans cette gamme, à moins que ce ne soit *avant* la mesure à partir de considérations aprioriques qui, plus souvent qu'autrement, risquent de faire référence à des fréquences relatives observées.

Ce jugement sur la mesure, c'est-à-dire celui à l'effet qu'elle offre un accès épistémique privilégié et qu'elle peut en conséquence être considérée objective, n'élimine cependant pas la possibilité de l'erreur. Loin de là. Et cela pour deux raisons. D'une part, il n'élimine pas d'emblée le doute sceptique avec lequel doit conjuguer toute connaissance. Demeure donc les difficultés liées à une prétendue interprétation universelle, ou à celles liées à l'indétermination de la traduction et de la référence, tel que discutée par Quine (1953, 1960 et 1992). On peut toujours en principe soutenir, par exemple, que l'occurrence « pile » ne sera pas interprétée de la même façon ni même assentie par tous les observateurs, ou même affirmer, selon une conception kantienne, que certains protocoles de mesure ne satisfont pas aux conditions formelles de possibilité d'une connaissance universellement partagée. Cependant, même une définition de la probabilité comme une « relation entre une personne et le

¹⁴ Dans le cas du dé, par exemple, la probabilité d'une occurrence est évidemment déterminée par rapport à la gamme {« pile », « face »}, et affirmer que cette probabilité est différente de $\frac{1}{2}$ revient soit à définir une mesure de probabilité inégale (par ex. le dé est biaisé), soit à présupposer une troisième possibilité (par ex. « tomber sur la frange »), auquel cas il faut redéfinir la gamme des possibles.

monde » (Lindley 2006 : 38) établit une distinction entre un discours sur le monde et le monde lui-même. Nul besoin donc de présumer que ce genre de relativisme puisse menacer l'aspect objectif de la mesure. D'autre part, si l'attribution d'une valeur de probabilité exige la détermination d'une gamme de possibles, elle ne peut par contre garantir l'accès épistémique à cette gamme. Autrement dit, on ne peut supposer que les connaissances sont toujours optimales afin de garantir la stabilité d'une valeur de probabilité. On peut bien attribuer des valeurs de probabilités différentes à différents moments, mais cela ne menace pas non plus l'objectivité (Handfield 2011).

Cet aspect objectif est toutefois insuffisant à caractériser la méthode d'attribution des probabilités. Il faut aussi faire appel à ces « inférences informelles ». Si la mesure permet d'établir des possibilités, et *in fine* une gamme de possibles¹⁵, l'interprétation des résultats de mesure et donc de la fréquence relative n'est pas encadrée clairement. Si, par exemple, un dé est lancé plusieurs fois et que la gamme {1, 2, 3, 4, 5, 6} est déterminée, rien n'indique *a priori* la valeur de probabilité de telle possibilité ou de tel type d'occurrence. Généralement, on infère, de séquences plus ou moins semblables pour chacune des possibilités, des valeurs de probabilités égales pour chacune de ces possibilités. Autrement dit, on infère l'équiprobabilité ou on attribue une mesure de probabilité uniforme. Mais cette inférence ne reçoit d'autres justification que le succès de la règle : elle permet de faire des prédictions relativement fiables qui ne sont pas contredites par l'expérience. L'absence de règles affecte aussi le fractionnement des unités individuelles d'une gamme de possibles, principalement pour des systèmes plus complexes qu'un dé par exemple. Un autre exemple est relié aux difficultés de la paramétrisation d'un continuum, comme l'illustre le « paradoxe de Bertrand »¹⁶. Il n'y a donc pas de règles claires pour le choix d'une mesure de probabilité appliquée aux systèmes réels. C'est aussi le cas dans le choix de ce qui peut faire partie des conditions initiales « acceptables ». Par exemple, des impulsions jugées *trop* semblables risquent de biaiser la distribution supposément uniforme des résultats. En termes plus techniques, si l'espace des phases (ou des états) ainsi que la dynamique du semi-groupe¹⁷ sont considérés objectifs, le choix d'une mesure théorique¹⁸ de probabilité, en revanche, peut sembler arbitraire et (donc ?)

¹⁵ Dans le cas d'événements très rares cependant, on peut objecter que la mesure (expérience) ne révélera cette possibilité qu'à condition de procéder à un très grand nombre d'essais. Il serait plus juste de dire, toutefois, que la mesure ne peut garantir l'occurrence de tels événements, mais il reste néanmoins possible qu'ils se produisent *et* qu'ils soient observés lors de la mesure.

¹⁶ Ce paradoxe met en évidence les limites du recours à l'intuition dans l'application du calcul des probabilités ainsi que la dépendance à la méthode de sélection des paramètres jugés pertinents et plus particulièrement de la variable aléatoire. Dans sa formulation originale par Joseph Bertrand, il consiste à choisir au hasard une corde d'un cercle donné et d'estimer la probabilité que celle-ci soit de longueur supérieure au côté du triangle équilatéral inscrit dans le cercle. Il en donnait alors trois réponses différentes (une chance sur deux, une sur trois et une sur quatre), toutes les trois apparemment valides.

¹⁷ Un semi-groupe est une structure algébrique consistant en un ensemble muni d'une loi de composition interne (une application) associative.

¹⁸ Une fonction associant une grandeur numérique à certains sous-ensembles d'un ensemble donné ; par ex. la mesure de Lebesgue.

« subjectif ». À défaut de règles claires pour encadrer ce genre d'inférence, une part irréductible d'arbitraire s'insère dans la méthode d'attribution de valeurs de probabilités.

L'appareil des probabilités au sein d'une théorie scientifique joue donc sur deux plans, ou modes, en inférant à partir de règles mathématiques, bref en calculant, d'une part, et en tentant d'identifier empiriquement des systèmes probabilistes et de justifier, empiriquement toujours, l'attribution de valeur de probabilités à ces systèmes. Le défi d'une interprétation des probabilités consiste ainsi à unifier ces deux plans ou modes impliqués dans l'appareil probabiliste. Mais est-ce conceptuellement possible ?

5.1.3.2 PROBABILITÉ EN PHYSIQUE

Afin de bien illustrer cette difficulté d'unification conceptuelle, je propose de donner quelques exemples, tirés de la physique, ce qui permettra par ailleurs de jeter les bases d'une interprétation originale des probabilités (Jodoin 2013c). Le résultat d'un tir de pièce de monnaie, par exemple, est considéré certain » et objectif parce qu'il permet d'une part, d'établir l'un des possibles, et d'autre part, de corroborer ou de corriger un énoncé probabiliste, et ce par une mesure, par l'observation. Bien qu'il puisse mener à la révision de certaines croyances, il est évident, toutefois, qu'il ne peut s'agir de falsification au sens logique du terme, car un énoncé probabiliste n'exclut rien d'observable. Cet aspect objectif de la mesure se manifeste aussi en physique. La mécanique statistique importe ainsi de la thermodynamique les mesures fixant la valeur des paramètres macroscopiques permettant de définir l'ensemble des micro-états possibles, c'est-à-dire compatibles avec ces mêmes mesures. En mécanique quantique, la mesure microscopique d'un système quantique est jugée certaine et permet de définir l'ensemble des micro-états possibles au sein d'un formalisme définissant un outil d'évaluation probabiliste polyvalent que l'on appelle le « vecteur d'état ».

Ainsi, dans ces différents exemples, ce qui est jugé certain et objectif dans la description d'un système est compatible avec un certain nombre d'informations sous-déterminant la description d'une autre composante de ce système, c'est-à-dire d'une indétermination au sens ontologique du terme ou d'une connaissance incomplète. Bien que cette autre composante, à défaut d'un meilleur qualificatif, puisse faire l'objet d'une sous-détermination, il n'est pas nécessaire ici d'exposer le préjugé déterministe. C'est-à-dire qu'une interprétation des probabilités n'implique pas une ontologie particulière : dans un monde ontologiquement déterministe, les probabilités renvoient nécessairement à des connaissances incomplètes, comme dans la conception classique laplacienne, tandis que dans un monde ontologiquement indéterministe, les probabilités peuvent autant renvoyer à des connaissances

exhaustive du monde qu'à l'incertitude de l'évaluateur¹⁹. Il ne semble pas nécessaire, cependant, que ce soit à une interprétation des probabilités d'apporter une réponse à cette question métaphysique. Il est d'ailleurs largement admis que cette question est empiriquement indécidable (Popper 1934 ; Bitbol 1997 ; Werndl 2009). L'interprétation propensionniste se rabat uniquement sur l'aspect objectif des probabilités et doit, afin d'éviter d'être taxée de trivialité où les probabilités ne prendraient que les valeurs 0 ou 1, adopter une ontologie indéterministe, incompatible au demeurant avec la mécanique statistique (Clark 2001).

La sous-détermination de cette autre composante peut néanmoins être caractérisée selon un rapport ou bien synchronique ou bien diachronique à ce qui est jugé « certain » et « objectif ». Dans l'exemple de la mécanique statistique, le micro-état est « synchroniquement sous-déterminé » parmi l'ensemble des micro-états possibles établis par la mécanique d'après le macro-état, lui-même déterminé par la mesure, car la mesure, à un certain moment et déterminant le macro-état, ne détermine pas exactement quel est, à ce même moment, le micro-état. Similairement, en mécanique quantique, le micro-état est « diachroniquement sous-déterminé » parmi l'ensemble des micro-états possibles, établis par le formalisme et des mesures préalables, car une mesure déterminant un micro-état, à un certain moment, ne détermine généralement pas quel sera, à un moment ultérieur, le résultat d'une autre mesure. La situation n'est pas différente à ce dernier cas pour le tir d'une pièce de monnaie où l'ensemble des états possibles, macroscopiques cette fois, est déterminé par des mesures préalables²⁰.

Ces exemples tirés de la physique permettent d'établir une distinction entre *probabilité synchronique* et *probabilité diachronique*. La terminologie est seulement différente mais les concepts sont les mêmes que ce que présente Leeds (2003) avec *probabilité d'être* (« *probability of being* »), où étant donné un système dans un macro-état particulier il y a une certaine probabilité qu'il soit dans un certain micro-état (« réalisant » ce macro-état), et *probabilité de devenir* (« *probability of becoming* ») où la spécification de l'état du système est cohérente avec des probabilités gouvernant son comportement futur. Albert (2000 : 80, n8) fait aussi cette distinction entre une règle de probabilité d'*incidences* de propriétés et une règle de probabilité de *séquences* de propriétés. Enfin, Loewer (2001, in Frigg 2008 : 674 ; aussi 2004 : 1116) semble aussi faire cette distinction lorsqu'il parle d'un « genre de chance dynamique » (« *a kind of dynamical chance* ») à propos d'un système macroscopique auquel on associe une distribution micro-canonique.

¹⁹ Par ailleurs, je doute que dans un monde absolument déterministe le concept d'*essai indépendant* fasse vraiment sens.

²⁰ Un examen sommaire de la pièce peut paraître suffisant pour établir cet ensemble des possibles (pile ou face), en sorte que la mesure ne serait pas nécessaire. Cependant, non seulement l'inférence allant des caractéristiques établies par cet examen sommaire (symétrie, etc.) aux états possibles ne va pas de soi, mais il semble bien que déjà cet examen constitue une mesure (au sens large) par l'observation, une relation mesureur-mesuré.

Prenons maintenant un exemple simple décrit formellement et similairement à ce que proposent ces deux théories physiques. Un dé présente six occurrences possibles $\{1, 2, 3, 4, 5, 6\}$. Cette gamme de possibles est établie par des mesures définissant un contexte expérimental particulier et elle est objective, dans le sens donné précédemment. Or, le résultat projeté d'un tir particulier est bien sûr compatible avec cette gamme de possibles, mais aussi avec un ensemble de conditions initiales, lesquelles demeurent inconnues et indéterminées par la mesure. Il est tenu pour acquis que si ces conditions initiales étaient déterminées, donc connues, alors la gamme de possibles se réduirait à une seule possibilité car ce système qu'est le dé est considéré être régi par des lois déterministes et objectives. Il est possible, par ailleurs, qu'il soit virtuellement impossible de déterminer parfaitement ces conditions initiales, mais là n'est pas la question. Cette gamme de possibles peut être associée à un espace des phases composé de six régions. Dans un cas déterministe comme celui-ci, ces régions correspondent aux évolutions temporelles possibles du système, de son état initial à son état final. Et comme cet état initial est inconnu, soit la position du dé et l'impulsion qui lui ait donnée, alors la région dans lequel se trouve le dé dans son état final est aussi inconnue. Autrement dit, l'information sur les conditions initiales est insuffisante pour en prédire le résultat. Et c'est ainsi que « s'introduit la notion de hasard » (Diu 1999 : 767). La mesure du dé dans son état final détermine une région de l'espace des phases définissant l'occurrence, soit le chiffre indiqué sur le dé, *et* un ensemble de conditions initiales compatibles avec cette mesure et l'évolution considérée causale et déterministe du système selon les lois de la mécanique classique. Mais la mesure du dé dans son état final ne détermine pas quelles étaient précisément ces conditions initiales, et c'est la raison pour laquelle seulement une région et non un point de l'espace des phases est déterminée.

Ce n'est pas le dénombrement de la totalité des conditions initiales qui permet d'attribuer un volume égal à chacune de ces régions de l'espace des phases, mais plutôt l'inférence impliquant un nombre d'occurrences à peu près égal pour chacun des résultats possibles *et* la sous-détermination des conditions initiales. Car dans un contexte déterministe, la détermination de ces conditions implique aussi celle de l'occurrence, du résultat final. Les résultats de mesure sont ainsi compatibles avec un certain nombre d'informations, comme un ensemble de conditions initiales, ici la position du dé et l'impulsion qui lui ait donnée, des lois mécaniques déterministes et aussi la façon dont est déterminé, mesuré, le résultat final. Dans cet exemple du dé, ce qui est objectif sont les mesures qui déterminent non seulement la gamme des possibles mais aussi les occurrences particulières parmi cette gamme. On pourrait y ajouter la détermination de l'évolution nomique du système selon les lois de la mécanique, mais il suffit ici de considérer le caractère déterministe de l'évolution associant deux régions de l'espace des phases. C'est-à-dire, chaque micro-état initial, qu'il soit empiriquement accessible ou non, est théoriquement associé à un micro-état final. La mesure et le calcul sont donc deux aspects importants,

considérés comme objectifs, dans l'application du calcul des probabilités au traitement des phénomènes physiques.

Ce qui peut être considéré comme étant subjectif, en revanche, n'est pas aussi simple. Car, dans un contexte déterministe, la sous-détermination des conditions initiales ne fait sens que par rapport à une supposée possibilité d'une connaissance exhaustive incluant ces conditions initiales. La sous-détermination ne peut alors que renvoyer à l'état de connaissance incomplet d'un sujet. Mais personne ne semble arguer que les jeux de dés constituent des exemples probants d'un monde indéterministe. Le caractère probabiliste apparaît ici dans cette incertitude quant aux conditions initiales, une sorte d'« aveuglement volontaire » pour que la notion de hasard puisse s'appliquer, même si en pratique la détermination de ces conditions initiales semble impossible à mettre en œuvre. Car si ces conditions initiales étaient connues, compte tenu des lois déterministes censées s'appliquer à un système comme le dé, l'état final serait tout aussi bien connu, déterminé, et la probabilité de l'état final compte tenu de l'état initial serait égale à l'unité. De plus, il demeure toujours possible d'introduire une probabilité conditionnelle à une information subséquente et qui modifie la valeur de probabilité initiale. Par exemple, la probabilité d'obtenir une certaine valeur lors d'un lancé de dé peut être conditionnelle à l'information selon laquelle le dé se trouve dans telle ou telle position à un millimètre de la table. Il est donc tentant d'apposer une grille de lecture subjectiviste où la probabilité n'est fonction que des connaissances d'un sujet susceptible de mettre à jour ses informations sur un système donné.

Pour bien illustrer l'absence de règles claires, formelles, comme dans le cas des règles mathématiques du calcul des probabilités, ce qui semble justifier l'étiquette « subjective » des probabilités, je propose un rapprochement avec le problème de l'induction. De fait, il est reconnu qu'il n'y a pas de règle claire encadrant le nombre d'observations nécessaire pour valider l'inférence partant d'énoncés singuliers à un énoncé général. De même, il n'y a pas de règle claire permettant de circonscrire une fréquence relative et de lui conférer une probabilité. Ce problème est persistant en sciences puisqu'un énoncé général voire universel, comme une loi scientifique, est souvent inféré à partir de l'observation de plusieurs cas particuliers. Une inférence déductive à partir d'un énoncé universel peut être rigoureusement valide mais la conclusion n'est vraie que si les prémisses sont vraies. Et comment s'assurer que les prémisses sont vraies ? Généralement, par inférence inductive, qui ne peut être rigoureusement valide à l'instar de sa contrepartie déductive. Il y a là un sérieux problème de circularité qui nous place aux limites des sciences empiriques. Cette circularité est aussi présente dans les inférences probabilistes. En effet, alors que des données empiriques sont nécessaires afin de construire une distribution de probabilité, cette dernière est nécessaire afin de déterminer le nombre suffisant de données empiriques. L'attribution d'une valeur de probabilité à un événement de même

que le choix d'une mesure de probabilité sont donc aux prises avec des difficultés du même ordre que celles reliées à l'induction.

Il convient maintenant de revenir aux exemples de la physique. En mécanique statistique, une mesure permet de déterminer l'état thermodynamique (macroscopique) du système définissant des paramètres devant être statistiquement corrélés avec les variables mécaniques. Une mesure ne détermine toutefois pas spécifiquement le micro-état du système, c'est-à-dire qu'elle ne peut spécifier un point représentatif dans l'espace des phases, mais plutôt un ensemble ou, si l'on veut, un grand nombre de points représentatifs. La mécanique statistique ne décrit pas dans quel micro-état *est* un système, mais plutôt quels sont les micro-états *possibles* ou quel *pourrait être* le micro-état s'il était empiriquement accessible. Ainsi, un ensemble de plusieurs micro-états est compatible avec un ensemble de paramètres définissant l'état thermodynamique, de sorte le micro-état est « synchroniquement sous-déterminé » par rapport à cet état thermodynamique, macroscopique. Comme, *ex hypothesi*, il n'y a pas de moyen empirique de déterminer quel est le micro-état du système, l'assignation d'un micro-état ou de plusieurs micro-états à un état macroscopique se fait de manière « orthogonale », soit par une méthode qui n'est pas définie par la théorie des probabilités, en l'occurrence par les équations de la mécanique et certaines hypothèses jugées raisonnables. Cet ensemble de micro-états définit donc une gamme de possibles qui est mise en rapport avec les paramètres définissant l'état macroscopique, mais « encadrant » aussi cette une gamme de possibles. Cette mise en rapport fait l'objet d'une attribution d'une valeur de probabilité fixée, généralement, par l'uniformisation de la gamme de possibles où chacune de ces possibilités a un « poids » égal, une équiprobabilité. Or, la justification de l'équiprobabilité est généralement considérée comme insatisfaisante. Ce serait donc là, par une hypothèse que « notre ignorance nous condamne à [...] adopter » (Poincaré 1902 : 203), que s'introduirait le caractère subjectif de la probabilité, comme « mesure d'un degré de croyance ». Mais, encore une fois, ne retenir que cette sous-détermination d'où découle notre ignorance revient à gommer un pan essentiel de la notion de probabilité en physique qu'est l'objectivation des résultats de mesure et de l'évolution nomique des systèmes pouvant faire l'objet de telles mesures.

En mécanique quantique, des mesures, dans un contexte donné permettant de définir des tests dits « quasi-identiques », permettent aussi d'établir une gamme des états possibles. Pour un système et un contexte donnés, la probabilité est déterminée de manière inférentielle par la *fréquence relative* de ces états possibles (on oublie trop souvent ce fait théorique). Ces informations, si l'on veut, sont incorporées au sein d'un formalisme définissant un « vecteur d'état » (ou « fonction d'onde »), comme outil de prédiction probabiliste, d'une part, et d'autre part, comme ce qui représente le mieux l'état du

système²¹. Ce vecteur d'état exprime la teneur prédictive d'un contexte expérimental donné et il est défini dans un espace abstrait ayant autant de dimensions que de faits élémentaires d'une certaine gamme de possibles expérimentaux. En termes simples, le calcul de la probabilité d'un événement particulier, d'un événement possible, suit un algorithme relativement simple de projection géométrique, où sa valeur est déterminée par le carré du module de la projection du vecteur d'état sur l'un des vecteurs de base de norme unitaire. Ainsi, la mesure implique un changement discontinu, passant du vecteur d'état distribué dans cet espace abstrait à l'un de ses vecteurs propres (qu'on appelle « effondrement de la fonction d'onde »). La mécanique quantique ne décrit pas dans quel micro-état *sera* un système, mais plutôt quels sont les micro-états *possibles* ou quel *pourrait être* le micro-état lorsqu'il sera empiriquement accessible par la mesure. Ainsi, un ensemble de résultats de mesure possibles est compatible avec un ensemble de paramètres définissant le vecteur d'état et le contexte expérimental, de sorte que l'état mesurable est « diachroniquement sous-déterminé » par rapport à ce vecteur d'état dans ce contexte donné. La situation n'est donc pas radicalement différente que celle du dé si l'on considère que la préparation d'un système quantique dans un contexte expérimental particulier ne permet pas de déterminer, de connaître entièrement les conditions initiales (mais on considère généralement que le vecteur d'état donne une connaissance « complète » du système).

Comme premier rapprochement entre ces deux théories, on peut constater qu'il semble impossible d'obtenir une mesure spécifiant un micro-état compatible avec une évolution temporelle, une séquence d'états. En d'autres mots, il semble impossible, à partir d'un micro-état mesuré et de lois dynamiques, de prédire le résultat d'une mesure à un temps ultérieur, soit parce que cet état est inaccessible, comme en mécanique statistique, soit parce que la description nomique de l'évolution temporelle du système ne permet pas de prédire univoquement le résultat mesure, comme c'est le cas en mécanique quantique. On peut donc y aller d'une amorce de définition. En mécanique statistique, d'une part, la probabilité est fonction de la compatibilité entre la mesure obtenue de l'état, qui est donc accessible et macroscopique, et l'ensemble des micro-états possibles, c'est-à-dire *calculés* à partir des lois de la mécanique classique, mais néanmoins sous-déterminés. D'autre part, en mécanique quantique, la probabilité est fonction de la compatibilité entre la mesure obtenue du micro-état, qui est donc accessible et partiellement macroscopique par son couplage à l'instrument de mesure, et l'ensemble des

²¹ Cette affirmation est lourde de conséquences philosophiques. Les débats concernant la question de ce que représente vraiment le vecteur d'état et de ce qu'est un état quantique sont nombreux et toujours d'actualité. Prétendre le contraire serait déjà prendre partie. En tout état de cause, le formalisme, avec sans doute « un peu d'interprétation », si tant est que tout formalisme est dénué de contenu sémantique, stipule que le vecteur d'état caractérise l'état du système quantique. Affirmer que ce vecteur d'état, représenté aussi par une « fonction d'onde », constitue une « description complète du système » ou la « somme de nos connaissances » est une tout autre histoire. Par ailleurs, les conséquences du théorème de Bell posent de sérieuses difficultés, comme la non-localité, au réalisme scientifique, dans sa mouture générique, sans toutefois le rejeter complètement.

micro-états possibles, suite à une certaine évolution temporelle, c'est-à-dire *calculés* à partir des lois de la mécanique ondulatoire, qui sous-déterminent aussi le résultat de la mesure. Les probabilités expriment donc le rapport de l'objectivation des résultats de mesures et d'une description nomique du système sous-déterminant, de manière ou bien synchronique ou bien diachronique, l'état du système parmi une gamme de possibles.

Cette sous-détermination peut être interprétée et justifiée de plusieurs façons. Il est tentant une fois de plus d'y appliquer une grille de lecture subjectiviste référant à des connaissances incomplètes. Mais cette lecture ne fait sens que dans la possibilité d'une exhaustivité principielle des connaissances. Cela ne va pas de soi en mécanique quantique. Puisqu'il s'agit d'un rapport, il ne va pas de soi non plus que le caractère subjectif suspendu à des connaissances incomplètes soit prédominant par rapport au caractère objectif de la mesure et de la description nomique. Aussi, même si le dénombrement des possibles, théoriquement ou empiriquement, semble objectif, leur mise en rapport quantitatif au sein du calcul des probabilités suit généralement des hypothèses qui peuvent paraître arbitraires et, partant, subjectives. Par exemple, le principe de raison insuffisante, qui devrait « guider notre ignorance », a été sévèrement critiqué et se prête à une lecture subjectiviste (Uffink 1995). Mais même une conception objectiviste comme l'interprétation fréquentiste a maille à partir dans la justification de l'inférence passant de l'observation d'une fréquence relative à la probabilité des occurrences. Le problème ne pourra être tranché empiriquement car aucun énoncé probabiliste n'est falsifiable au sens fort, logique du terme. Cet aspect arbitraire et informel, « subjectif » donc, un terme quelque peu malheureux puisqu'il gomme partiellement la richesse du concept de probabilité, semble inévitable. Mais ce concept est subjectif (mieux : arbitraire) autant que peut l'être le concept d'induction, en ce qu'il ne spécifie pas le nombre nécessaire de cas pour soutenir l'inférence à l'énoncé général. Il est donc aux prises avec des problèmes épistémologiques aussi vastes et importants que celui, par exemple, de « théorie scientifique », ou de « loi » (à quel moment passe-t-on d'une généralisation à une loi ?). C'est pourquoi il serait mal avisé de le réduire à des « degrés de croyance ».

5.1.4 Remarques conclusives

La preuve de l'importance des probabilités en sciences n'est plus à faire. Les conséquences philosophiques de cet usage accru depuis plus d'un siècle ne sont toutefois pas aussi claires. Il ne faut s'en étonner si, comme on l'a vu, l'interprétation des probabilités fait intervenir tout un réseau de problèmes philosophiques particulièrement importants pour la philosophie des sciences et, pourrait-on ajouter, de la philosophie tout court. Par conséquent, mais aussi en raison de, les interprétations sont

nombreuses. Le point départ de l'argumentation visant à établir une (autre) interprétation des probabilités est constitué ici d'un constat et d'un argument. Le constat est celui à l'effet que la pléthore d'interprétations présente différentes prises de position par rapport au schème dualiste de la connaissance, le couple objectif-sujetif, qui forme aussi, bien entendu, un critère de classifications de ces interprétations parmi d'autres. Mais de ce constat ne découle l'obligation de trancher, au sens d'écarter l'une des options, ce schème dualiste (du moins ce fût le point de vue que j'ai adopté). L'argument concerne la distinction entre calcul des probabilités, sur le plan formel, et attribution de probabilités, sur le plan empirique. Si le premier semble établi et, dans une certaine mesure, prémuni contre l'ambiguïté, le second doit trouver sa définition opératoire au terme d'un exigeant et néanmoins rigoureux examen épistémologique. Car une probabilité ne s'observe pas, ni se déduit d'une description du monde. J'ai aussi avancé que l'inférence visant à établir dans les faits une valeur de probabilité est de même « nature » que celle visant à établir la nécessité d'un phénomène, ce qui n'est pas sans rappeler la fameuse critique humienne.

Partant, donc, de l'affirmation que la probabilité ne peut être identifiée à la fréquence relative, qui n'en constitue pas moins, dans certains cas à définir, une justification valable, et qu'elle ne peut, de plus, être déduite d'arguments aprioriques, il s'ensuit que son introduction au sein de sciences empiriques ne peut faire l'économie d'un élément d'arbitraire, de subjectivité. Un élément aussi arbitraire, cependant, que peut l'être l'inférence inductive passant de l'observation de cas répétés à un énoncé universel, à une loi au sens fort du terme. Mais les probabilités, dans leur attribution de valeur dans les faits, procède aussi, et j'ose dire surtout, de l'objectivation des résultats de mesures et de descriptions nomiques. C'est pourquoi tout ce qui est possible n'est pas nécessairement probable. C'est-à-dire que tout phénomène ne rencontre pas nécessairement les conditions d'objectivation encadrant l'attribution de probabilité. C'est aussi pourquoi elles n'offrent pas de réponses permettant de surmonter le schème dualiste de la connaissance, et que leurs interprétation s'inscrivent dans un débat beaucoup plus large qu'il n'est généralement admis. Cette approche des probabilités ne désigne pas de gagnants entre la théorie subjectiviste et la théorie objectiviste ou propensionniste, mais propose plutôt deux perdants.

5.2 Mécanique statistique

La mécanique statistique est née, tel qu'indiqué précédemment, d'une tension entre la mécanique et la thermodynamique. Son objet principal est de dériver des lois macroscopiques à partir de lois microscopiques sous-jacentes. Elle est fondée sur l'utilisation des probabilités en raison de l'énormité du nombre de molécules (de l'ordre de la constante d'Avogadro, soit environ 6×10^{23}) associé à la quantité de systèmes microscopiques. Elle a cette particularité de fournir des méthodes considérées comme rationnelles pour traiter des comportements de systèmes mécaniques où les connaissances de leur état sont moindres que celles qui seraient théoriquement possibles d'obtenir. Il s'agit donc de dresser un portrait du comportement global des éléments d'un tout. Selon Penrose (1970), le point de départ de la mécanique statistique est le fait que l'on ne peut déterminer l'état dynamique (microscopique) d'un système physique macroscopique. Bref, il faut faire beaucoup avec peu.

Alors comment la mécanique statistique s'y prend-elle pour rendre compte de l'entropie thermodynamique ? En réalité, on peut répondre : de plusieurs façons. En effet, la tension initiale entre la mécanique et la thermodynamique semble être l'obstacle premier dans l'uniformisation des méthodes de conceptualisation de l'entropie statistique en ce qu'il faut combler un fossé théorique qui paraît laisser une trop grande liberté au théoricien. Comme le soutient Sklar (1993 et 1999), il y a plusieurs choses en mécanique statistique pouvant être associées à l'entropie thermodynamique, et de plusieurs façons. Bien entendu, ce n'est pas tant la liberté du théoricien qui pose problème que le manque de guide ou d'information permettant de caractériser l'entropie dans le paradigme de la mécanique statistique. De plus, l'essentiel de la base empirique est constitué des phénomènes thermodynamiques et en ce sens l'objectif d'une définition de l'entropie statistique revient en grande partie à une reformulation réductionniste de la définition en thermodynamique. Malgré les nombreuses affirmations voulant que l'entreprise réductionniste de la mécanique statistique soit un succès, il y a lieu d'être dubitatif face à la prétention que l'entropie ait été ramenée ou réduite à la dynamique.

Il y a essentiellement deux approches à la mécanique statistique et donc deux formalismes, quoique semblables, et donc au moins deux types de définitions de l'entropie statistique. Il s'agit de l'approche de Boltzmann et celle de Gibbs, et des définitions qui s'y rattachent. Outre la définition de l'entropie thermodynamique (Chapitre 4), la taxonomie de l'entropie statistique comporte quatre définitions de l'entropie.

5.2.1 Formalisme

Le formalisme de la mécanique statistique se présente essentiellement en deux versions, celle de Boltzmann et celle de Gibbs. Ces deux formalismes sont semblables, puisqu'ils comportent ces deux ingrédients principaux que sont la mécanique classique et le calcul des probabilités, mais aussi l'hypothèse fondamentale à l'effet qu'un système physique est représenté par un système mécanique avec un très grand nombre de degrés de liberté dans un espace euclidien. Mais ils sont aussi, bien entendu, différents puisque l'un se focalise sur les systèmes individuels, tandis que l'autre sur les ensembles de systèmes (explicités plus loin). La différence de ces approches peut mener à des conséquences épistémologiques divergentes impactant profondément la définition et l'interprétation du concept d'entropie.

D'abord, un système physique est représenté par un système mécanique avec un très grand nombre de degrés de liberté, appelé aussi *sous-système*. Ce dernier est défini dans l'*espace des phases* Γ qui est un espace conceptuel de dimensions $2fN$ constitué d'axes orthogonaux des variables canoniques, soit un construit de l'espace de configuration formé par les coordonnées de position q_1, \dots, q_N et de l'espace des moments formé par les moments conjugués p_1, \dots, p_N . La dimension de l'espace des phases pour un système tridimensionnel est donc $2 \times 3N = 6N$. En mécanique classique, qui constitue évidemment la théorie de base de la mécanique statistique, son « fond ontologique » en quelque sorte, l'état microscopique instantané d'un système peut être considéré comme un *point représentatif*, x , dans l'espace des phases. L'espace des phases pour un système vu comme un tout (par ex. un gaz) est appelé espace- γ (ou Γ_γ), tandis que l'espace des phases pour un élément quelconque (par ex. une molécule) est appelé espace- μ (ou Γ_μ). Utilisant un espace- γ approprié, l'état de n'importe quel système peut être complètement spécifié par la position d'un point représentatif dans cet espace. Pour un système à N particules, l'espace- γ est ainsi un espace à $6N$ dimensions, de sorte $x_\gamma = \{q_1, \dots, q_N, p_1, \dots, p_N\}$ représente un vecteur dans l'espace- γ et réfère au *micro-état*. L'ensemble qu'est l'espace- γ est le produit cartésien de N copies de l'espace des phases espace- μ à six dimensions d'une particule.

L'espace des phases définit ainsi l'ensemble des états possibles du système, et plus précisément du sous-système caractérisé par un état dynamique, un état microscopique. Toutefois, une telle précision dans la représentation du sous-système par x dans Γ_γ n'est ni nécessaire ni même désirable. D'abord, il semble clair qu'en pratique les moyens expérimentaux dont nous disposons soient insuffisants pour une telle caractérisation et qu'en théorie, la mécanique quantique, une certaine indétermination demeure inévitable. Ensuite, il est possible et même raisonnable de croire que le petit nombre de paramètres nécessaires pour rendre compte des phénomènes thermodynamiques puisse être légitimement représenté par des quantités « globales », comme des moyennes sur cet ensemble. D'autant plus qu'il

n'est pas si évident, il semble même à première vue complètement erroné, que l'ensemble des particules puisse être représenté par un continuum d'états mécaniques. Enfin, comme dans l'approche gibbsienne, l'attention peut être portée sur les multiples variations de faible intensité que les particules peuvent globalement subir.

Ainsi, *par hypothèse*, l'état dynamique, microscopique, d'un système relativement grand, ceux dont les propriétés sont susceptibles d'être mesurées et dont traite la thermodynamique, n'est *pas* accessible à la mesure. La motivation de cette hypothèse est généralement considérée d'ordre pratique en ce que nos moyens expérimentaux ne garantissent pas cette accessibilité. C'est pourquoi l'espace des phases peut être constitué d'une collection de régions élémentaires égales, appelées aussi *cellules*, de dimension $h = \delta p \delta q$, où h est une constante à déterminer, et donc de volume h^{3N} . Avec le recul de la physique contemporaine, il peut être tentant d'opter pour la constante de Planck tirée de la mécanique quantique. En effet, le théorème d'indétermination formulé d'abord par Heisenberg postule l'impossibilité de principe, de déterminer, de mesurer simultanément la position, avec une précision δq , et le moment, avec une précision δp , d'une particule sans la contrainte $\delta q \delta p \geq \hbar$, où \hbar est la constante de Planck. Cependant, bien que cet emploi soit incohérent au sein d'une description classique, il peut tout de même constituer une unité de mesure. Donc, en bref, il ne fait « aucun sens » (Gallavotti 1999 : 4) de tenter de déterminer le micro-état avec plus de précision. En ce sens, l'idée de la description d'un état (total) précis est dans tous les cas un « concept limite abstrait » (Tolman 1938 : 1).

Comme toute quantité physique est fonction des variables dynamiques du système, c'est-à-dire de son espace des phases, l'évolution de l'état du système doit aussi être celle des variables dynamiques. L'hypothèse atomique sur laquelle s'appuie la mécanique statistique implique une application (mathématique) qui transforme les cellules de l'espace des phases mutuellement, décrivant ainsi la dynamique du système. La loi d'évolution, déterminée par les lois de la mécanique, impose ainsi à chaque région et à chaque cellule de l'espace des phases des valeurs à certains paramètres (énergie cinétique, énergie potentielle, énergie totale) selon un principe déterministe. Ces quantités sont fonctions des quantités de mouvement et des positions des particules dans l'état microscopique correspondant au centre de la région ou cellule considérée. Cette région doit être la plus petite afin de se rapprocher le plus possible d'un état complètement spécifié, représenté par un point, mais, tel que mentionné, il n'est pas accessible, et c'est pourquoi l'on considère de petites cellules. Les paramètres associés aux cellules sont donc ceux déterminés par son centre, et bien qu'un autre point que le centre dans la cellule déterminent d'autres valeurs pour ces paramètres, la différence doit être *inobservable*, autrement les cellules ne seraient pas, *a contrario ex hypothesi*, les plus petites à être observables.

Non seulement les expériences, les mesures des quantités physiques ne permettent pas de spécifier le micro-état, mais aussi elle ne rapportent pas les valeurs instantanées de ces mêmes quantités physiques. Plutôt, chaque mesure ou observation doit durer un certain temps qui peut paraître très court à l'échelle humaine, mais qui est très long à l'échelle microscopique où les états subissent de nombreux changements, en raison, principalement, des nombreuses collisions moléculaires. Par conséquent, ce qui est expérimentalement accessible se limite à des *moyennes temporelles* – $f^* = \lim[T \rightarrow \infty] 1/T \int_0^T f(t) dt$, où f représente une quantité physique donnée – sur de longues périodes. De fait, la mesure d'une quantité thermodynamique (macroscopique) est supposée durer plus longtemps que le temps t_c requis pour la disparition des corrélations moléculaires (par exemple, les interactions à la collision), mais tout de même moins longtemps que le temps t_m maximum au cours duquel la quantité macroscopique ne change pas : $t_c < t < t_m$. Or, ce temps maximum peut être considéré comme *infini* puisqu'à l'équilibre le système est considéré ne plus subir de changement observable. Malheureusement, ces moyennes temporelles sont calculées par intégration des équations dynamiques où il est absolument vain de tenter d'obtenir une solution exacte pour de si grands ensembles. C'est pourquoi la *moyenne en phase* – $\langle f \rangle = \int f(x) \rho(x) dx$ – qui représente la valeur attendue ou prévue, du système est utilisée. Le problème, appelé *problème ergodique*, est donc de savoir si, ou à quelles conditions, la moyenne temporelle égale la moyenne en phase. Cette prétendue égalité, qui demeure contestée, correspond à la condition d'*ergodicité*.

Ainsi, en adoptant condition d'ergodicité et en définissant un volume \mathbf{v} de l'espace des phases formé à partir de petits intervalles de coordonnées, il est possible de définir une durée Δt où le sous-système se trouve dans \mathbf{v} et représentant une portion d'un temps T suffisamment long, de sorte que

$$w = \lim[T \rightarrow \infty] \Delta t / T.$$

Cette quantité peut alors être considérée, avec tout ce que cela comporte de polémique épistémologique, comme représentant la *probabilité* que le sous-système soit dans un volume \mathbf{v} donné à un instant arbitraire. Il va donc de soi que cette probabilité est déterminée par des moyennes temporelles. Mais, nonobstant la condition d'ergodicité, on peut tenter d'établir la probabilité qu'un système soit dans un certain état par la configuration des molécules dans les différentes cellules de l'espace des phases. Car, pour des propriétés observationnelles (macroscopiques) données, il y a plusieurs « occupations » de cellules différentes compatibles avec ces observations. C'est pourquoi il est préférable de se concentrer sur ce qu'on peut appeler la « condition » du système, c'est-à-dire sur ce qui caractérise le système de manière compatible avec les informations disponibles, sans toutefois le spécifier de manière exhaustive du point de vue microscopique. Il y aura donc, pour N molécules et en négligeant les permutations, un total de $\Omega = N! / (n_1! n_2! n_3! \dots n_i! \dots)$ cellules différentes dans l'espace- γ correspondant à la condition du système, déterminée par le nombre de molécules n_i dans les différentes

cellules i de l'espace- μ . La probabilité (fréquentiste) P de trouver le système dans une condition est alors proportionnelle à cette quantité (Tolman 1938 : 78) :

$$P = \mathcal{N}! / (n_1! n_2! n_3! \dots n_i! \dots) \times \text{constante} \propto \Omega.$$

En mécanique statistique, tel que mentionné, il y a essentiellement deux approches historiquement et épistémologiquement différentes, celle de Boltzmann et celle de Gibbs. Dans l'approche de Boltzmann l'état d'équilibre est une propriété observable des systèmes *individuels* et la probabilité d'une transition de cet état à un autre est nulle. Cette approche suppose que la séquence markovienne des états observationnels inclut des états déterministes, soit les états d'équilibre (Penrose 1970 : 157). Dans l'approche de Gibbs, en revanche, l'état d'équilibre est une propriété observable d'un *ensemble* de systèmes et les probabilités pour tout état observable sont indépendantes du temps, du moins si on les définit pour des régions suffisamment « grossières » de l'espace des phases. Tandis que dans l'approche de Boltzmann, les concepts d'équilibre et d'entropie sont définis à partir d'un macro-état, lui-même caractérisé par des variables macroscopiques, l'approche de Gibbs les définit à partir d'une distribution, d'une fonction de densité de probabilité de l'espace des phases.

On peut distinguer deux types de questions épistémologiques ayant trait à la formulation de l'entropie statistique en général : des questions quant à l'*état* décrit par cette quantité et celles quant à son *évolution*. Les premières renvoient principalement à l'interprétation des entités statistiques, comme les densités de probabilités, qui interviennent dans la définition de l'entropie. En ce sens, il s'agit bien souvent de questions de *référence* (à quoi réfèrent ces quantités ?), à savoir principalement si la définition réfère à des systèmes individuels ou à des ensembles de systèmes. Il y a aussi des questions importantes quant à la *pertinence* (la définition de l'entropie est-elle une définition de la physique, est-elle utile et dans quel contexte, etc. ?), car bien souvent il est question de systèmes infinis. Et il y a bien sûr les questions d'interprétation faisant intervenir des métaphores comme celles du désordre ou de l'étalement (« *spreading* »). Le second type de questions épistémologiques, qui ne sont évidemment pas étrangères aux premières, concerne les hypothèses permettant de retrouver une fonction ayant un comportement dans le temps *analogue* à celui de l'entropie thermodynamique, qui est censée augmenter dans le temps. En ce sens, il peut y avoir un certain décalage entre les réponses apportées aux questions sur l'état et celles sur l'évolution ; ainsi, les premières questions peuvent recevoir des réponses satisfaisantes sans toutefois répondre aux exigences des secondes ; c'est le cas, par exemple, des définitions statistiques impliquant une invariance temporelle. Dans une telle situation où les objectifs généraux pour une définition statistique ne sont pas atteints, il est légitime de remettre en question non seulement la *méthode* employée que le *projet* lui-même, tel que défini par la mécanique statistique.

5.2.2 Entropie de Boltzmann

L'entropie de Boltzmann se présente principalement sous trois variétés qui découlent toutes du *théorème H*, lequel s'appuie sur la validité de l'*équation de transport de Boltzmann* et sur celle de l'hypothèse du *Stosszahlansatz* (Chapitre 3). Malheureusement, ce théorème fait face à de nombreux problèmes, dont le principal est sans doute celui de justifier un comportement irréversible à partir de lois réversibles. Au surplus, les autres hypothèses sur lesquelles s'appuient les différentes versions de l'entropie de Boltzmann sont aussi très problématiques. L'idée de base des entropies de Boltzmann, qui semble évidente mais qui ne l'est pas au regard de l'approche de Gibbs, est que celles-ci doivent refléter le comportement de l'entropie thermodynamique comme propriété d'un système individuel et elles doivent augmenter dans le temps avant d'atteindre un maximum à l'équilibre.

5.2.2.1 ENTROPIE CONTINUE DE BOLTZMANN

Le premier type d'entropie statistique selon Boltzmann est l'*entropie continue de Boltzmann*, ou à *grenage fin*, qui est une conséquence directe du théorème *H*. Elle est définie telle que

$$S_{B,f} \equiv -kNH[f_i],$$

où k est une constante et f_i est une distribution ou densité définie sur $\Gamma\mu$ (Boltzmann 1872, in Brush 2003 : 291 ; Wherl 1991 : 119 ; Frigg & Werndl 2011 : 123). Tel que présenté précédemment, dans son (1872), Boltzmann est en mesure de prouver qu'une quantité $H[f_i]$ décroît de manière monotone dans le temps. Or, pour une distribution stationnaire f^* d'un gaz idéal monoatomique, comme la distribution de type maxwellien, il obtient $H[f^*]$, une quantité correspondant à l'entropie thermodynamique si l'on a défini l'entropie telle que $S_{B,f}$.

Il va donc de soi que la validité de $S_{B,f}$ dépend de celle du théorème *H* : si $S_{B,f}$ est associée ou identifiée correctement à S_T alors le théorème *H* pourrait fournir une « interprétation microscopique » (Boltzmann 1872, in Brush 2003 : 263) ou une « justification » (Frigg & Werndl 2011 : 123) du second principe de la thermodynamique. Mais il est loin d'être clair s'il s'agit d'une preuve que l'entropie doit augmenter ou d'une simple interprétation du second principe, bref le statut épistémologique de ces résultats demeure largement ambigu. La raison principale de l'identification du concept d'entropie thermodynamique à celui d'entropie statistique est que la fonction associée à cette dernière est définie avec des variables dynamiques et qu'elle se comporte *comme* la fonction d'état de la thermodynamique. Il s'agit donc d'un raisonnement par analogie formelle (Chapitres 2 et 6). De plus, avec ce théorème Boltzmann prouve à la fois plus et moins que nécessaire : plus, en ce sens que le second principe ne parle ni d'entropie pour les états de non-équilibre, ni d'accroissement monotone ; et moins, car

Boltzmann ne considère pas l'accroissement de l'entropie pour des processus généraux adiabatiques, puisqu'il traite plutôt de gaz isolés de leur environnement, une condition plus forte. Ainsi, ni le théorème H ni, donc, la définition de l'entropie continue ne sont exempts de problèmes théoriques et épistémologiques, que je traiterai dans la discussion qui suit et qui reviendront dans celles vouées aux autres définitions de l'entropie.

Les questions quant à l'état décrit par l'entropie continue touchent surtout l'interprétation de $f(x, t)$. Comme distribution cette quantité est censée représenter la fraction du nombre total de molécules, ou le nombre relatif de molécules, du système se trouvant dans une région particulière de l'espace des phases. La question se pose alors de savoir si cette interprétation en termes « mécaniques » exprime oui ou non une probabilité. Cette lecture stricte, comme fraction ou nombre relatif, n'implique pas nécessairement une notion de probabilité. Mais Uffink (2007 : 967) y voit une indication chez Boltzmann que les probabilités constituent un langage approprié pour traiter des problèmes mécaniques de la théorie des gaz. Mais cela est vrai seulement si cette fraction ou ce nombre relatif est justement interprété, avec la figure d'autorité qu'incarne Boltzmann, comme une probabilité. Il est en effet possible d'appliquer une grille de lecture en termes probabilistes selon une interprétation fréquentiste : si on pige des molécules du gaz de manière aléatoire, avec réinsertion, alors la distribution $f(x, t)$ représente la fréquence relative de molécules dont l'état est défini par une certaine région de l'espace des phases. D'autres interprétations sont évidemment possibles (discutées plus loin).

Il faut toutefois remarquer que cette distribution est continue alors qu'elle doit être, en fait, discrète, car elle est constituée d'un nombre (N) fini de points dans l'espace- μ à six dimensions, l'espace des phases des molécules individuelles. Deux arguments peuvent justifier cet emploi (Zeh 1989 : 35s). D'une part, on peut faire appel à la limite thermodynamique, lorsque le nombre de molécules tend vers l'infini ($N \rightarrow \infty$), où la différence entre une fonction continue et une fonction discrète devient négligeable, et où les temps de récurrence (retour à un état arbitrairement proche de l'état initial) deviennent infinis. Mais nul besoin d'avoir recours à l'infini pour contrer l'objection de la récurrence, car les temps de récurrence sont de l'ordre de plusieurs milliards d'années pour les systèmes observables (Chapitre 3). Cette idéalisation d'une limite asymptotique présente certes des avantages mathématiques mais son application à la description de systèmes physiques « réels » ne va pas de soi. Les systèmes infinis ne sont *pas* les systèmes que l'on cherche à expliquer. Or, *en théorie*, les transitions de phases, comme le passage de l'eau de l'état liquide à l'état gazeux, ne se produisent que dans les systèmes infinis, que si la limite thermodynamique est appliquée. C'est-à-dire que les singularités présentes dans le traitement thermodynamique de ces phénomènes apparaissent dans le traitement qu'offre la mécanique statistique à condition que certains paramètres des fonctions à dérivées partielles disparaissent et qu'une discontinuité apparaisse, ce qui est mathématiquement possible à la limite

thermodynamique. Mais évidemment, *en pratique*, on en voit à tous les jours. L'apparente nécessité de cette idéalisation semble donc en faire un critère essentiel à la distinction entre les domaines descriptifs microscopique et macroscopique. Comme la propriété observable décrite par la thermodynamique ne peut être récupérée qu'à l'aide d'une telle idéalisation, il s'agirait d'un cas d'*émergence* (Liu 1999 et 2001 ; voir aussi les Chapitres 2 et 6). Mais il n'est pas inévitable de calquer en mécanique statistique le traitement thermodynamique, de prendre cette dernière « trop au sérieux » (Callender 2001), bien que ce genre d'idéalisation de systèmes infinis et de limite asymptotique puisse être nécessaire à notre compréhension de ces phénomènes de transition de phase (Battermann 2002 et 2005).

D'autre part, la justification d'une distribution continue peut se faire en accolant une « incertitude » à l'état dynamique des particules, c'est-à-dire en considérant non pas des points dans l'espace des phases mais plutôt de petits éléments de volume dans l'espace- γ contenant une infinité de points. Ces éléments de volume permettent de définir une distribution continue telle que $f(x, t)$. Mais définir de la sorte l'état du système mène à des incongruités dans la description de son *évolution*. Car l'ensemble de points de l'élément de volume ne peut être établi de manière univoque à partir de f définie dans l'espace- μ car celle-ci ne contient pas l'information quant aux corrélations entre molécules (Zeh 1989 ; Ridderbos 2002). La distribution continue définie dans l'espace- μ ne définit qu'un macro-état et ne spécifie donc pas l'état dynamique microscopique. Il s'ensuit que la dérivée par rapport au temps de cette distribution qui permettrait de décrire l'évolution du système ne peut être obtenue à partir de l'évolution de l'élément de volume. Il faut donc procéder autrement pour déterminer l'évolution temporelle de la distribution.

Par ailleurs, les questions quant à l'évolution se réfèrent directement ou indirectement à la base sur laquelle s'appuient les versions statistiques, soit le théorème H . Ce théorème peut être interprété comme une « mesure de l'écart du micro-état d'un système avec un certain état d'équilibre » (Tolman 1938 : 134), lequel peut aussi être vu comme un « attracteur universel »²² (Prigogine & Stengers 1979 : 277), en ce qu'il ne dépend pas du mécanisme d'interaction entre molécules, ou de l'étalement de la distribution des vitesses des molécules. Son principal problème concerne le paradoxe à l'effet que le modèle mécanique complètement réversible d'un gaz parvienne à expliquer les processus irréversibles (j'y reviendrai ; voir aussi les Chapitres 3 et 4). Ce théorème est basé sur un certains nombres d'hypothèses, qui rendent sa portée plus restreinte. Il s'appuie d'abord sur le modèle d'un gaz à N molécules, d'un seul type et sphériques, ce qui exclut les mélanges. Aussi, le gaz est si dilué ou raréfié que seules les collisions binaires sont prises en comptes, il ne peut donc pas s'appliquer aux liquides ni

²² Puisqu'une dynamique réversible préservant la mesure ne possède *pas* d'attracteur (Lavis 2005), et comme la mécanique classique décrit une telle dynamique, il y a lieu de douter de la possibilité même de prouver le théorème H .

aux solides. Mais, enfin, le gaz ne peut être trop raréfié, car si le nombre de molécules est trop petit, l'hypothèse de continuité, permettant de passer d'une sommation à une intégrale, est injustifiée.

L'hypothèse cruciale, pourtant considérée comme anodine par Boltzmann (1872) puisqu'il ne lui a offert aucune justification, est celle concernant le nombre de collisions moléculaires, soit le *Stoßzahlansatz* (Chapitre 3). En substance, cette hypothèse stipule que le nombre de collisions, pour deux groupes de molécules en mouvement l'un vers l'autre, est déterminé par le produit de la somme des volumes cylindriques, formés par le déplacement des molécules sphériques du premier groupe, et du nombre de molécules du second groupe par unité de volume. Ainsi, selon le *Stoßzahlansatz*, il n'y aurait aucune corrélation entre molécules, autrement dit les collisions seraient aléatoires. Alors pourquoi le *Stoßzahlansatz* est-il une hypothèse d'asymétrie ? Parce que, bien que les molécules ne soient pas corrélées *avant* la collision, supposément parce qu'il n'y aurait pas de raison de croire autrement, elles sont censées devenir corrélées *après* ou *en conséquence* de leurs collisions. Le raisonnement appuyant la première partie de cette hypothèse s'appuie sur une autre hypothèse d'équiprobabilité que les Ehrenfest (1912 : 6 ; n25) appellent « hypothèse sur le chaos moléculaire », comme quoi le nombre de molécules par unité de volume des molécules entrant en collision est partout le même dans l'espace considéré. Mais cette hypothèse, qui peut être interprétée, ou non, d'après le principe de raison insuffisante, comme quoi nous n'aurions par de raison de croire qu'il y a corrélation entre paires de particules sur le point de collisionner, n'est *pas* une hypothèse dynamique. Elle provient « d'ailleurs » et introduit subrepticement un élément d'irréversibilité, étranger à la mécanique classique. (Pour des discussions plus étendues à ce propos, voir Zeh 1989 ; Price 1996 ; Brown, Myrvold & Uffink 2009.)

La stratégie de Boltzmann pour déterminer l'évolution temporelle de la distribution f consiste à exprimer sa dérivée partielle par une équation offrant un bilan en termes de taux de transition pour un type de collision $(\mathbf{v}_1')(\mathbf{v}_2') \rightarrow (\mathbf{v}_1)(\mathbf{v}_2)$ et son inverse $(\mathbf{v}_1)(\mathbf{v}_2) \rightarrow (\mathbf{v}_1')(\mathbf{v}_2')$. Le résultat est une équation différentio-intégrale, l'*équation de Boltzmann* (voir Boltzmann 1872, in Brush 2003 : 279 ; Cercignani 1998 : 260 ; Uffink 2007 : 964 ; Müller 2007 : 95 ; Chapitre 3). L'application de cette équation à la quantité $H[f]$ montre que celle-ci croît dans le temps et qu'elle s'accorde, à une constante près, avec l'expression thermodynamique de l'entropie pour les gaz parfaits. Ce résultat n'est pas banal. Mais sa réception mitigée provient sans doute de la prétention de son auteur d'avoir offert une preuve rigoureuse, analytique et générale au second principe de la thermodynamique. D'une part, l'hypothèse asymétrique qu'est le *Stoßzahlansatz* mérite de plus amples arguments, car elle semble ne reposer que sur un caractère d'évidence autoproclamée. Le problème, qui se manifeste dans l'objection de l'irréversibilité, est que deux descriptions du monde tenues pour fondamentales mais pourtant logiquement incompatibles, telles quelles, soit celle de la mécanique classique et celle la thermodynamique phénoménologique, sont censées s'unifier au sein d'un raisonnement où la première

expliquerait la seconde. D'autre part, les suppositions par rapport aux corrélations entre particules sont problématiques car elles peuvent devenir déterminantes dans des systèmes autres que des gaz parfaits.

Comme pour la majorité des formulations de l'entropie, l'expression de $S_{B,f}$ découle de deux propriétés fondamentales qui sont tirées de la formulation thermodynamique puis attribuées à sa formulation statistique : (i) la *non-décroissance* pour les processus adiabatiques, ce qui entraîne son lot de problèmes, comme on l'a vu, et (ii) l'*additivité* où l'entropie de deux systèmes est la somme de deux entropies individuelles. Pour ce qui est de (i), la preuve découle du théorème H , tel que discuté. En termes simples, l'application de l'équation à la quantité $H[f_i]$ serait donner une quantité $(a - b) \log(b/a)$, laquelle est toujours négative si a et b sont des réels positifs (voir Brush 1976 : 238 ; Müller 2007 : 98). L'additivité en revanche est nécessaire si l'on veut être cohérent avec l'expérience et avec une définition de la température faisant intervenir la dérivée partielle de l'entropie. Cependant, en mécanique statistique, il est impossible de satisfaire complètement (i) et (ii) à moins d'employer la limite thermodynamique, donc d'avoir recours à l'infini (Penrose 1970 : 156 ; discuté plus loin).

5.2.2.2 ENTROPIE COMBINATOIRE DE BOLTZMANN

Le second type d'entropie statistique selon Boltzmann est l'*entropie combinatoire de Boltzmann*, ou *principe de Boltzmann* selon l'expression d'Einstein (1905), et il est notoirement exprimé par

$$S_{B,c} = k \ln \Omega$$

où k est une constante et Ω (parfois W pour *Wahrscheinlichkeit*) est le nombre de complexions ou d'arrangements (Wherl 1991 : 120 ; Müller 2007 : 1000 ; Frigg & Werndl 2011 : 124). Mais d'où vient-il ? En fait, dans la littérature, ce principe est rarement dérivé ou justifié mais est plutôt seulement défini ou stipulé. Il y a donc des lacunes à combler : un besoin d'une preuve, d'une justification de ce principe.

Si l'on suppose que le théorème H est bien valide, ce qui ne va pas de soi, il est possible d'obtenir l'entropie combinatoire de Boltzmann via l'*argument combinatoire* (voir Cercignani 1998 : 121s ; Müller & Weiss 2005 : 27s ; Müller 2007 : 99-103 ; Uffink 2007 : 975-6 ; Frigg 2008 : 13s). L'argument repose sur la discrétisation ou la partition de ce qui est appelé l'espace- μ , Γ_μ , soit l'espace des phases à six dimensions d'une particule ou molécule particulière. La partition de l'espace- μ en m cellules disjointes ω_i est une procédure de *grenage grossier* (« *coarse graining* ») : $\mu = \omega_1 \cup \dots \cup \omega_m$, où le *micro-état* à *grenage grossier* (aussi appelé un « arrangement », ou « *Komplexion* » par Boltzmann) d'un tout comme un gaz est une collection de N points dans Γ_μ , un point pour chaque particule. Une fonction de distribution discrète est alors une collection d'entiers positifs n_i et correspond au nombre de particules dont l'état est dans la cellule ω_i , ce qui définit un *macro-état* (« distribution d'état », « *Zustandsverteilung* ») : $\mathcal{Z} := (n_1, \dots,$

n_m). Ainsi, si les molécules ne sont *pas* interchangeables, donc discernables, la quantité Ω est définie comme

$$\Omega(\mathcal{Z}) \equiv \mathcal{M}/n_1!, \dots, n_m!$$

et donne le nombre d'arrangements compatibles avec une distribution \mathcal{Z} donnée, un macro-état donné. La relation entre micro- et macro-état n'est donc pas unique car des micro-états différents sont obtenus par la permutation des molécules : il y a plusieurs micro-états compatibles avec un macro-état donné. C'est pourquoi on considère qu'un arrangement, un micro-état, contient « plus d'information » (Albert 2000 : 43 ; Frigg & Werndl 2011 : 124) qu'une distribution, qu'un macro-état.

L'argument combinatoire comporte évidemment plusieurs hypothèses. L'expression de $S_{B,c}$ n'est pas évidente, à première vue, compte tenu que le point de départ est encore le théorème *H*. La simple affirmation à l'effet que la version statistique devrait être une fonction générale ayant comme argument la quantité Ω ne va pas de soi. Ceci dit, *si* l'on tient pour acquis que cela doit être le cas, *si* cette quantité apparaît importante et interprétable au point d'en faire un point de départ, ou presque, à la définition de l'entropie statistique, alors il est vrai, compte tenu de certains critères comme l'additivité, que la fonction la plus simple pouvant satisfaire tous ces desiderata est la fonction logarithmique. Mais il est évident que la justification pour la quantité Ω demeure lacunaire, voire absente. Il faut en substance obtenir cette quantité à partir d'une fonction de la forme $f \ln f$. La stratégie de Boltzmann (1872), qu'il considérait comme un « artifice », consiste alors à discrétiser, à partitionner l'espace des phases de manière à attribuer à chacune des particules une cellule issue de cette partition et à laquelle est attribuée une valeur d'énergie, elle-même attribuée aux particules (voir Dugdale 1996 ; Müller 2007).

La première de ces hypothèses, donc, et sans doute la plus importante, concerne la partition de l'espace des phases afin de dénombrer les micro-états. Le problème avec cette stratégie de partition est double. D'une part, il faut justifier l'entreprise elle-même car il n'y a rien en mécanique classique qui permette une discrétisation de l'espace des phases et un tel dénombrement. En effet, les informations que l'on peut obtenir à propos d'un système, par une mesure définissant les paramètres macroscopiques comme l'énergie totale du système, est compatible avec une *infinité continue* de microconditions, entendues comme des ensembles de variables dynamiques pour chacune des particules. Car la totalité des microconditions du système newtonien possède invariablement la cardinalité d'un continuum. D'autre part, il faut justifier le choix de la partition, le choix d'une unité de dénombrement définissant la forme et la dimension des cellules. Généralement, chaque cellule est définie cubique et d'égal volume dans l'espace des phases (l'espace- μ). Une partition est ainsi une collection d'ensembles disjoints et mesurables, au sens mathématique du terme, couvrant tout l'espace des phases compatible avec le

principe de conservation de l'énergie. Tel que mentionné, il est tentant de garnir l'ontologie classique d'éléments tirés de la mécanique quantique qui offre, justement, une ontologie pour ainsi dire « discrétisée », faisant appel au « quantum d'action » représenté par la constante de Planck. Mais il faut bien voir qu'il y a conflit métaphysique et méthodologique. Car, sans pour autant suggérer qu'il y ait une correspondance univoque entre une théorie et « son » ontologie, le « quantum d'action » est incompatible avec l'ontologie classique de la mécanique. De plus, son emploi dans l'argument combinatoire suggère que ces deux théories peuvent être mise sur le même plan de manière isolée, sans qu'il soit nécessaire de revoir l'ensemble des hypothèses qui étaient, jusque là, puisées uniquement dans l'une d'entre elles. En l'absence de critère clair, donc, le choix de la partition peut paraître arbitraire et ambigu, et toute chose étant égale par ailleurs, l'ambiguïté est à éviter. Cela étant dit, le résultat est pratiquement le même pour une latitude considérable de choix de découpage ou partition.

Les autres hypothèses importantes de l'argument combinatoire sont les suivantes (voir Müller 2007 : 100-3). D'abord, le nombre maximal de particules pouvant occuper chaque cellule est le même pour chacune des cellules. Les objections que l'on peut adresser à cette hypothèse proviennent de la mécanique quantique et soulèvent donc, une fois de plus, un lot de questions sur les relations inter-théoriques. En bref, la discrétisation de l'espace des phases par le quantum d'action défini par la constante de Planck implique celle de certaines propriétés comme l'énergie, mais ne mène pas nécessairement à discrétiser uniformément les propriétés dynamiques des particules ou encore à discrétiser uniformément le nombre de particules. Ensuite, la façon de procéder au dénombrement des micro-états possibles est basée sur l'hypothèse que la permutation de particules identiques mène à des micro-états différents. Toutefois, et cela amène aussi plusieurs questions d'ordre métaphysique, la mécanique quantique semble plaider en faveur du principe de l'identité des indiscernables (van Fraassen 1991 ; Bitbol 1997 ; Ladyman & Ross 2007 ; Riggs 2009), de sorte que les différentes permutations de particules considérées identiques n'en constituent en fait qu'une seule. C'est pourquoi ce passage dans l'argument combinatoire, qui demande l'ajout, qui paraît arbitraire, du facteur M , est en réalité moins anodin qu'il n'y paraît puisqu'il est généralement justifié par l'hypothèse que les particules sont, contrairement à ce qu'en dit le point de vue classique, indistinguables. On constate donc que ce problème renvoie au concept de l'haecécité²³ sans toutefois y apporter une réponse (Huggett 1999 ; Albert 2000). Enfin, le passage de la fonction de la forme $f \ln f$ à la quantité Ω demande l'approximation de Stirling qui n'est valide que si le nombre de particules dans les cellules est

²³ La doctrine ontologique associée à ce concept défend la thèse selon laquelle deux mondes ne diffèrent uniquement que par la permutation des positions de particules autrement identiques sont, néanmoins, différents. Ce problème mériterait de plus amples discussions mais je ne m'y attarderai pas davantage.

très supérieur à l'unité. Malheureusement, ces hypothèses sont « rejetées par la physique moderne » (Müller 2007 : 103).

Les fondations théoriques de l'entropie combinatoire de Boltzmann ($S_{B,c}$) paraissent donc peu convaincantes. Pourtant elle demeure utilisée et même prisée. On peut déplorer que les hypothèses menant à la définition de la quantité Ω , et donc de $S_{B,c}$, sont rarement exposées et encore moins discutées. La raison en est sans doute que, malgré les doutes que l'on peut avoir quant aux hypothèses censées la soutenir, le *résultat* est interprétable, du point de vue l'analyse combinatoire, comme le nombre d'arrangements possibles compatibles avec une distribution donnée de N particules. La situation est donc similaire à celui d'un jeu de cartes où l'analyse combinatoire peut aussi s'appliquer. Mais l'analogie demande plus de précision car la détermination de la gamme des possibles d'un système mécanique n'est pas la même que celle d'un jeu de cartes, ne serait-ce que parce que dans ce dernier cas les états possibles sont déterminés soit par la définition même de ce qu'est un jeu de cartes, soit par l'observation. Le fractionnement des unités individuelles de la gamme des possibles (section 5.1.3) ne semble pas aussi clair. Donc, l'entropie combinatoire aurait-elle une signification allant bien au-delà de celle issue de la théorie cinétique des gaz, doit-elle être « vraie sans égard à ces fondations » comme le suggère Müller (2007 : 102) ? La réponse réside dans l'interprétation qu'on en fait. Car, au juste, quel état, le cas échéant, $S_{B,c}$ décrit-elle ? Elle est souvent décrite comme exprimant une « compatibilité » entre le nombre d'arrangements et une distribution donnée, entre une quantité de micro-états pour un macro-état donné. De manière plus ou moins équivalente, mais qui mérite un examen réservé pour plus tard (sections 3.2.3 et 3.4.4.), on admet que $S_{B,c}$ décrit l'« accessibilité » (Capek & Sheehan 2005 ; Frigg & Werndl 2011) des micro-états pour un macro-état, ou encore une relation entre « l'information spécifiant un micro-état et la complexité du système » (Greven, Keller & Warnecke 2010 : 2).

Étant donné la relation un-plusieurs entre un macro-état et ses micro-états associés, où le nombre de micro-états est calculé à partir de contraintes macroscopiques telles l'énergie du système, son volume et le nombre de molécules, une description possible est celle de la *réalisabilité multiple* (Chapitres 2). En tout état de cause, il est question de potentialité, ce qui s'exprime par des conditionnels : le système se trouve dans l'un de états possibles dénombrés de manière combinatoire, et s'il était dans tel micro-état particulier, il aurait tel caractéristique observable. On y définit donc une gamme de possibles pour un état sous-déterminé. Il n'est donc pas étonnant que les probabilités entrent en jeu, de même que le couple subjectif-objectif (section 5.1.3). Ainsi, une interprétation possible est une interprétation épistémique ou inférentielle à l'effet que $S_{B,c}$ serait une mesure de ce qui est possible (formellement ? pour nous ?) d'*inférer* à propos d'un arrangement d'un système à partir de sa distribution. Ou encore, elle serait une mesure de notre « indifférence dans les estimés de plausibilité » (Maes 2010 : 253).

L'interprétation de l'entropie combinatoire, qui en fait sa valeur si tant est que ses fondations sont précaires, a donc beaucoup à voir avec la notion de probabilité, mais aussi avec un principe stipulant l'évolution temporelle de cette entropie. Une interprétation de l'évolution de $S_{B,c}$ est celle que Müller (2007 : 101) appelle un peu pompeusement la « stratégie de la nature » : en raison du mouvement irrégulier des éléments du systèmes, ces sous-systèmes que sont les molécules, et de l'hypothèse voulant que chaque arrangement soit équiprobable, il est « éminemment raisonnable » que le système évolue vers une distribution présentant plus d'arrangements compatibles et, éventuellement, toute chose étant égale par ailleurs, vers la distribution présentant *le plus* d'arrangements compatibles. L'idée générale semble donc que le système évolue *d'un état improbable à un état plus probable*. Cette affirmation semble, intuitivement du moins, raisonnable, puisque, toute chose étant égale par ailleurs, parmi tous les états possibles pour tous les moments possibles, la majorité de ces états doivent être probables, *ex hypothesi*, et si ce n'est pas le cas, alors pas pour très longtemps. Mais elle mérite un examen plus attentif (prochaines sections). Quoiqu'intéressante à bien des points de vue, le problème avec cette interprétation est qu'elle semble contredire l'affirmation à l'effet que le système évolue en fonction de la dynamique de ses éléments. Car c'est bien de cela qu'il s'agissait au départ. La grille de lecture probabiliste est « orthogonale » à celle de la mécanique classique. Cette dernière a bien été choisie en raison de la confiance qu'on lui porte pour constituer ce « fond ontologique » des phénomènes observés que l'on cherche à expliquer. Comme le souligne Uffink (2007 : 979), les états n'évoluent pas vers d'autres états simplement parce qu'il y en aurait *plus* de ces derniers.

5.2.2.3 ENTROPIE MACROSCOPIQUE DE BOLTZMANN

Le troisième type d'entropie statistique selon Boltzmann est l'*entropie macroscopique de Boltzmann* qui se définit sur l'espace- γ ($\Gamma\gamma$) comme

$$S_{B,m}(M) \equiv k \ln \mu(\Gamma_M)$$

où Γ_M est un sous-ensemble de l'espace des phases d'énergie constante ($\Gamma_M \subset \Gamma \equiv \Omega_E$) représentant un état macroscopique M constitué de points en phase représentatifs qui sont macroscopiquement similaires à l'état microscopique x , et μ réfère au volume, donné par la projection invariante par rapport au temps de la mesure de Liouville (Lebesgue) sur la surface d'énergie (Lebowitz 1993 et 1999 ; Goldstein 2001 ; Lavis 2005 ; Frigg & Werndl 2011). Ainsi, chaque micro-état x de ce volume correspond à M , mais à un macro-état donné correspondent plusieurs micro-états. En ce sens, $S_{B,m}$ mesure la portion de l'espace- γ du système qui est occupée par les micro-états correspondant à l'état macroscopique M .

Cette définition est donc très semblable à l'entropie combinatoire (section 5.2.2.2). En tirant profit de l'extension mathématique de la notion intuitive de mesure, celle de Lebesgue, elle attribue une valeur de « volume » à une région de l'espace- γ plutôt qu'un nombre. Elle évite ainsi le problème du dénombrement de micro-états sur un continuum. Tel qu'indiqué plus tôt, l'espace- γ est le produit cartésien de N copies de l'espace- μ , en sorte que le choix d'une partition sur l'espace- μ induit une partition de l'espace- γ . Par conséquent, à chaque distribution « découpant » l'espace- μ correspond une région bien définie de l'espace- γ . Une question cruciale pour des raisons tant théoriques qu'épistémologiques revient donc à justifier ce découpage, cette partition de l'espace des phases. On a vu, d'un côté, que celle de l'espace- μ revenait à définir des « cellules » qui doivent être ni trop grandes ni trop petites. D'un autre côté, il n'y a pas de procédure canonique, consensuelle, pour découper l'espace- γ afin de déterminer clairement une région de l'espace des phases correspondant à un état macroscopique M . Une proposition consiste à découper l'espace des phases en fonction de la précision de nos moyens d'observation. Comment rendre explicite une telle proposition est toutefois difficile. La critique quant à son caractère arbitraire et ambigu revient donc ici. Que cette ambiguïté ait peu de conséquences pratiques fait peu, à ce niveau de la discussion, pour étouffer cette critique. C'est pourquoi on peut proposer, presque par dépit, que la région définie par le produit du nombre d'arrangements (Ω) et du volume de l'ensemble des cellules de cette distribution *est* la région associée à M (par ex. Frigg & Werndl 2011).

Quant à l'interprétation de l'entropie macroscopique, elle fait face aux mêmes enjeux que l'entropie combinatoire, principalement en ce qui concerne la notion de probabilité. Ainsi, la proposition précédente identifiant deux régions de l'espace des phases revient à définir, par la mesure constante du volume de la cellule, une équiprobabilité pour les micro-états. Et comme précédemment, d'aucuns suggèrent une interprétation épistémique ou inférentielle comme quoi l'entropie macroscopique mesurerait ce qu'on peut inférer sur la position du point représentatif dans la région associée à l'état macroscopique : plus grande est l'entropie macroscopique, plus grande est la région dans laquelle peut se trouver le micro-état. Pour se placer sur le plan de la métaphore et de l'expérience de pensée, la situation est similaire au cas d'un dé à plusieurs milliards de côtés, représentant les micro-états, dont plusieurs d'entre eux seraient peints d'une couleur particulière et reconnaissable, ce qui correspondrait aux états macroscopiques, les seuls étant empiriquement accessibles, observables. Alors, affirmer, suite à une observation, que le dé est en position, par exemple, « bleu », ne révèle pas sur quel côté est le dé, ne dit pas quel est son micro-état spécifique, même si cet état macroscopique peut recevoir une mesure de son « volume » et la valeur d'entropie correspondante. Il semble donc naturel, dans l'attente d'une justification, d'associer le volume en phase à la probabilité d'un état macroscopique (voir les prochaines sections).

5.2.3 Entropie de Gibbs

La mécanique statistique de Gibbs repose sur des hypothèses différentes de celles de Boltzmann. En effet, elle utilise ce qui est convenu d'appeler la *méthodologie des ensembles*, qui lui permet de traiter les caractéristiques des composants d'un système comme des variables indépendantes et de calculer des moyennes sur ces variables. Il a été vu qu'un *ensemble de Gibbs* est une collection infinie, indénombrable de systèmes de même structure qu'un système à l'étude, mais distribués à l'intérieur d'une gamme d'états possibles. Les éléments de cet ensemble n'interagissent pas entre eux, mais sont plutôt indépendants, car il s'agit d'une collection imaginaire. L'état d'un ensemble est donc spécifié par une distribution $\rho(x\gamma, t)$ exprimant les fractions des systèmes se trouvant dans un intervalle donné, $\{x\gamma, x\gamma + dx\gamma\}$, de l'espace des phases, l'espace- γ . Ainsi, une propriété comme l'entropie est définie en fonction de $\rho(x\gamma, t)$ et son évolution dans le temps est déterminée, en principe, à partir de l'évolution dans le temps de $\rho(x\gamma, t)$.

L'entropie de Gibbs à grenage fin $S_{G,f}$ est définie comme une fonctionnelle de fonctions de densité de probabilité ρ sur l'espace des phases :

$$S_{G,f} \equiv S[\rho] = -k \int \rho(x) \ln \rho(x) dx$$

(Gibbs 1902 ; Tolman 1938 ; Wherl 1991 ; Frigg & Werndl 2011). On peut désormais dire que l'approche de Gibbs fait partie de l'orthodoxie de la mécanique statistique. Cette définition de l'entropie, que Gibbs appelait plutôt « index de probabilité », repose sur des *analogies* entre des ensembles statistiques de molécules et certains résultats de la thermodynamique. La première de ces analogies consiste à associer l'équilibre statistique à l'équilibre thermique de sorte que l'hamiltonien d'un système composé équivaut à la somme des hamiltoniens des systèmes le composant (parce que l'hamiltonien d'interaction peut alors être considéré comme « infiniment petit »). Puisque la densité de distribution, ou de probabilité, $\rho(x\gamma, t)$, peut représenter l'état d'un ensemble de systèmes, l'équilibre thermodynamique est représenté naturellement par une densité qui est indépendante du temps. Ainsi, selon cette ligne de pensée, alors qu'à l'équilibre thermique entre deux corps correspond l'absence de changement dans le temps dans l'un ou l'autre de ces corps, l'équilibre statistique correspond à une situation où la probabilité de trouver certains points en phase dans certaines régions de l'espace des phases ainsi que les valeurs moyennes de l'ensemble sont aussi indépendantes du temps.

La seconde analogie consiste à reproduire l'« équation fondamentale » de la thermodynamique dans laquelle l'entropie apparaît. Gibbs fournit deux analogues à l'entropie, l'un est appelé l'*entropie de volume*, dont l'ensemble est le volume d'une région délimitée par l'hypersurface d'énergie constante dans l'espace des phases, et l'autre l'*entropie de surface*, dont l'ensemble est le volume d'une coquille d'épaisseur

infinitésimale d'énergie constante dans l'espace des phases, en quelque sorte l'« enveloppe » du volume précédent. Gibbs démontre que ces ensembles sont équivalents pour les grands ensembles. Mais il n'en reste pas moins qu'une sous-détermination demeure, qui ne s'efface, ou du moins elle devient en un certain sens négligeable, qu'à la limite thermodynamique, lorsque le nombre de molécules tend vers l'infini. De plus, bien que la littérature ait privilégié l'entropie de surface, il semble qu'il y ait des raisons théoriques de favoriser l'entropie de volume, car elle est un invariant adiabatique, à l'instar de l'entropie thermodynamique (Campisi 2005).

Les questions épistémologiques touchant l'état que décrit l'entropie de Gibbs concernent, à l'instar de l'entropie de Boltzmann, l'interprétation de la représentation statistique qu'est ici la distribution de probabilité $\rho(x)$. Cette distribution définie sur l'espace des phases est l'analogue mécanique statistique d'un état thermodynamique. Il a été vu que l'approche gibbsienne fait appel à la notion d'ensemble, soit une collection infinie de systèmes de même structure macroscopique qu'un système à l'étude mais distribués à l'intérieur d'une gamme d'états possibles des configurations microscopiques, des états dynamiques. L'état du système est donc représenté par une fonction continue définissant une gamme infinie de possibles et, plus spécifiquement, exprimant les fractions des systèmes se trouvant dans un intervalle donné de l'espace des phases $\{x\gamma, x\gamma + dx\gamma\}$.

Ainsi, l'entropie de Gibbs, $S_{G,f}$, est une fonction d'une distribution de probabilité, non pas d'une fonction de l'espace des phases. En ce sens, elle fait fi de ce « fond ontologique » de la mécanique, d'un monde constitué de particules régies par des lois du mouvement. Et Gibbs parle justement d'analogie et ne propose d'aucune manière une ontologie particulière (Chapitre 3). Ce faisant, par contre, on peut s'interroger si son entropie a ce qu'il faut pour décrire l'entropie thermodynamique. Et la réponse est « non » selon plusieurs, dont Albert (2000) et Goldstein (2001). Au surplus, il n'y aurait *aucune* fonction en phase χ , telle que $S_G[\rho\theta] = \langle \chi \rangle_T$, pour toute température T . Elle ne peut donc être interprétée comme la moyenne de « quelque propriété des membres individuels de l'ensemble » (Uffink 2007 : 995). Il y aurait donc un problème de référence, car si $S_{G,f}$ n'est *pas* une fonction de l'espace des phases, alors sa dépendance ou simplement sa relation à la dynamique microscopique gouvernée par les équations newtoniennes ou hamiltoniennes est remise en question. Il en va de même pour le projet même de la mécanique statistique, du moins dans sa conception initiale. Car l'idée que les propriétés macroscopiques thermodynamiques dépendent d'une certaine façon de la dynamique microscopique des molécules est essentielle au projet de la mécanique statistique. Par contre, l'idée que ces propriétés dépendent d'autres choses, de distributions de probabilité sur des ensembles infinis ou quoique ce soit d'autres, relève d'un autre projet.

Dans cette approche, la notion d'invariance temporelle est d'abord appliquée sur une distribution d'un ensemble caractérisant l'état d'équilibre, de sorte que $\partial\rho/\partial t = 0$. Autrement dit, l'équilibre y est défini comme la situation où la distribution n'est pas une fonction explicite du temps. Cependant, les conditions dynamiques sont insuffisantes à préciser une distribution à l'équilibre unique et par conséquent une infinité de distributions peuvent satisfaire ces conditions. C'est pourquoi le postulat d'équiprobabilité est généralement introduit afin de sélectionner une distribution unique temporellement invariante. Ce postulat stipule que, pour un système avec une certaine énergie, la distribution de probabilité est la distribution uniforme, appelée distribution microcanonique, sur cette région de l'espace des phases constituée de tous les micro-états compatible avec cette valeur d'énergie. Tous les autres micro-états se voient attribués une probabilité nulle. Cette idée de dépendance, qu'il faut préciser, entre les deux niveaux, l'un macroscopique et l'autre microscopique, est primordiale. L'objectif de la mécanique statistique est justement de « connecter », un terme faible laissant la marge nécessaire à la diversité des approches, deux théories en apparence incompatibles.

Quant à l'évolution, il y a le problème, voire le « paradoxe » (Wherl 1978), de l'*invariance temporelle* ou *invariance dynamique* : en raison du théorème de Liouville, la mesure de Lebesgue sur l'hypersurface d'énergie est invariante pour un flot hamiltonien et en conséquence l'entropie de Gibbs est une constante dynamique (Penrose 1970 ; Wherl 1978 et 1991 ; Goldstein 2001 ; Ridderbos 2002 ; Lavis 2005). Elle est donc indépendante du temps (c'est-à-dire que $dS_{G,f}/dt = 0$). Conséquemment, il est impossible pour une distribution d'ensemble de non-équilibre de changer, d'évoluer vers la distribution d'ensemble microcanonique caractérisant une situation d'équilibre. Elle ne permettrait donc pas de « comprendre » (Lebowitz 2000 : 11), et encore moins d'expliquer, l'évolution temporelle des systèmes macroscopiques. En effet, s'il n'y a pas d'évolution de la quantité censée représenter l'état d'un système, alors cette quantité et la formule qui la définit ne peuvent expliquer l'évolution de ce système. C'est précisément ce problème, pour ne pas dire cet échec, qui est à la base de l'inférence voulant que l'irréversibilité présente au sein du principe entropique *émerge*, que l'entropie thermodynamique soit une *propriété émergente* (Chapitre 2). Une solution, proposée par Gibbs, est celle du « *grenage grossier* », qui revient à partitionner, à construire des « cellules » sur l'espace- γ , ce qui rappelle bien sûr la procédure de Boltzmann, qui est plutôt sur l'espace- μ .

Cependant, il n'est pas si évident qu'il n'y paraît que l'invariance dynamique de $S_{G,f}$ pose un problème sérieux à sa candidature comme équivalent microscopique de l'entropie thermodynamique (S_T). En effet, d'une part, si le principe entropique stipule uniquement l'impossibilité d'une décroissance de l'entropie, alors une entropie constante répond trivialement à ce critère. Évidemment, on voit mal comment il pourrait expliquer le changement. D'autre part, l'entropie thermodynamique n'est (mathématiquement) définie qu'à l'équilibre et, en ce sens, trouver une fonction définie à l'équilibre,

dont les valeurs futures aussi à l'équilibre seraient plus élevées, serait suffisant. Or, l'approche gibbsienne n'a pas de problème à identifier une telle fonction et donc à fournir un équivalent microscopique de l'entropie thermodynamique pour les états d'équilibre. Une extension aux états de non-équilibre serait donc un atout intéressant mais non nécessaire. Mais le problème est ailleurs : étant donné le passage d'un état d'équilibre à un autre état d'équilibre, comme l'état macroscopique a changé, l'ensemble de Gibbs décrivant cet état doit changer aussi. Or, il ne semble pas justifié, ce ne serait pas « juste » (« *fair* », Sklar 1993 : 54), de simplement choisir un autre ensemble pour décrire le système à un temps ultérieur. On exige en effet une certaine continuité, établie ici par les équations de la mécanique classique, entre les états initial et final pour qu'on puisse bien parler d'évolution du système. Ainsi, l'affirmation à l'effet que la mécanique classique gouverne l'évolution du système serait fallacieuse.

Par conséquent, cette rupture d'avec la dynamique microscopique mène à une reconsidération, d'une part, de la possibilité d'une *réduction* de l'entropie thermodynamique, voire même du projet de la mécanique statistique puisqu'elle vise à fournir une traduction mécanique des quantités thermodynamiques, et, d'autre part, de la nature de la *référence* de $S_{G,f}$, qui semble perdre son assise microscopique. Car le caractère d'irréversibilité de la tendance à l'équilibre des systèmes thermodynamiques, dont on rend compte par le principe entropique stipulant que l'entropie d'un tel système augmente avant d'atteindre un maximum à l'équilibre (Chapitre 4), ne peut être expliqué par une quantité physique constante dans le temps. On constate ainsi, une fois de plus, que les problèmes concernant l'*état* de l'entropie touche de près à ceux concernant son *évolution*.

Ce problème devient plus manifeste dans le cas d'un analogue d'une quantité physique, puisque la justification d'une bonne analogie repose justement sur un comportement, une évolution similaire. Si la fonction des équations dynamiques ne se comporte pas comme celle de la thermodynamique, sur quelle base peut-on s'appuyer pour dire que cette fonction, parmi une infinité rappelons-le, et Gibbs en était fort conscient, est la « bonne » ? À la rigueur, le problème de la référence peut être court-circuité par le fait que $S_{G,f}$ n'est qu'un analogue, que Gibbs appelait d'ailleurs « index de probabilité », qu'elle possède la forme mais pas le contenu (bien que cette distinction mériterait plus d'explication) de l'entropie thermodynamique. Mais le problème de l'invariance temporelle n'est pas pour autant réglé. L'essentiel de la base justificative pour la connexion (un terme moins fort que « réduction ») entre ces deux définitions de l'entropie est constitué, d'une part, de l'analogie formelle, ce qui veut dire en termes simples que la forme de leur équation mathématique est à peu près la même, et d'autre part, des conséquences prédictives de $S_{G,f}$. Plus précisément, la connexion devient manifeste dans un cas spécial, lorsque la distribution canonique est appliquée à $S_{G,f}$, car on obtient une équation qui s'assimile à l'équation de l'entropie pour un gaz parfait.

Est-ce que tout cela fait sens ? Est-ce que la considération d'une infinité de systèmes semblables, mais mutuellement indépendants, et empiriquement inaccessibles, du moins en ce qui les distingue, à quelque chose à voir avec ce qui est raisonnablement recherché pour expliquer les systèmes thermodynamiques que nous observons ? N'est-ce pas « pure folie » (« *sheer madness* », Albert 2000 : 70) ? Même si le discours à l'effet que l'approche gibbsienne « explique » certains phénomènes thermodynamiques semble erroné ou peu convaincant, plusieurs raisonnements visant à le faire basent leurs inférences sur le *calcul* de l'entropie de Gibbs. Il est difficile à dire si le formalisme gibbsien, malgré les objections interprétatives que l'on peut lui adresser, se prête mieux, justement, au traitement mathématique permettant de retrouver certaines propriétés thermodynamique. L'approche adoptée pour l'interprétation des probabilités ne rejette pas d'emblée l'approche gibbsienne. En insistant sur la sous-détermination de l'état microscopique, le concept d'entropie de Gibbs sous-tend alors la notion de potentialité du système auquel le concept s'applique. De sorte qu'elle décrit ce qui est possible pour le système sans statuer sur son état actuel, lequel serait, par impossible, une sorte de superposition d'une infinité de sous-systèmes²⁴.

Un dernier point avant de terminer, pour l'instant, avec l'entropie de Gibbs. Il s'agit de ce qu'on appelle le « paradoxe de Gibbs ». Supposons deux gaz identiques, chacun ayant la même énergie U , volume V et nombre de particules N , dans des contenants adjacents, et supposons que l'entropie de chaque gaz soit $S(U, V, N)$, de sorte que l'entropie du système total soit $S_{\text{tot}} = 2S(U, V, N)$, ce qui répond à l'exigence d'extensivité ou d'additivité. Si une valve, par exemple, est ouverte de manière à faire la connexion entre les deux contenants, l'entropie S_{tot} devrait demeurer inchangée, tel que $S_{\text{tot}} = S(2U, 2V, 2N) = 2S(U, V, N)$. Toutefois, un calcul avec la définition *générique* de l'entropie, laquelle n'est pas extensive, prévoit que le simple fait d'ouvrir la valve conduit à une diminution d'entropie. Puisque, dans le cas de deux gaz identiques, le processus est considéré réversible, cela contredit le principe entropique. Une autre façon de formuler le problème est celle de Penrose (1970 : 171). Soit un mélange classique de N_1 particules de masse m_1 et N_2 particules de masse m_2 . Si $m_1 \neq m_2$, alors l'entropie d'un état observationnel est calculée à partir du facteur $N!$; et si $m_1 = m_2$, alors l'entropie de cet état observationnel est calculée avec $N = N_1 + N_2$. Ainsi, l'entropie de cet état observationnel pour $m_1 = m_2$ excède l'entropie du même état observationnel pour $m_1 \neq m_2$ d'une quantité $k \ln(N! / N_1! N_2!)$. Le paradoxe repose dans cette dépendance apparemment artificielle et discontinue de l'entropie par rapport à un paramètre continu comme $m_1 - m_2$.

²⁴ La distinction me paraît similaire à celle entre « infini actuel » et « infini potentiel ». C'est-à-dire que l'entropie de Gibbs peut être interprétée comme la gamme, infinie certes, des états potentiels du système, mais le système n'est *pas* dans une sorte de superposition infinie de micro-états.

La solution généralement acceptée consiste à répudier l'entropie spécifique et de choisir l'entropie *spécifique*, qui est obtenue en divisant la première par le nombre de permutations possibles, ce qui la réduit d'un facteur $1/M!$ par rapport à l'entropie générique. La justification de ce choix repose sur l'hypothèse que les particules sont indistinguables, donc qu'une simple permutation de particules ne mène pas à un nouvel arrangement (French & Rickles 2003 ; Swendsen 2008). Cette hypothèse peut bien sûr recevoir plusieurs justifications, physiques ou métaphysiques, comme celle invoquant le principe d'indétermination d'Heisenberg, selon lequel des particules supposément identiques ne peuvent être distinguées une fois qu'elles ont collisionné. Penrose (1970) soutient pour sa part que le passage, discontinu, d'une quantité petite mais non nulle de $m_1 - m_2$ à une quantité nulle n'est pas bénigne, car si l'influence dynamique est minime, négligeable, l'influence en revanche sur les propriétés *observables* peut être profonde, comme celle d'observer les concentrations de deux gaz différents. Ainsi, la justification se trouverait dans le changement de propriétés observationnelles qui doit être reflété dans un changement d'entropie. Pour Uffink (2007 : 999), dans l'approche spécifique, le mélange de substances, identiques ou non, mène à une augmentation d'entropie et il n'y a donc pas de paradoxe puisqu'il n'y a pas de changement abrupt de l'entropie, de sorte que la division par le facteur $M!$ ne permet pas de *résoudre* le paradoxe de Gibbs, mais plutôt il *évite*. Pour Ainsworth (2012) enfin, l'application de ce facteur, donnant ainsi une nouvelle définition de l'entropie de Boltzmann, implique que sa valeur, pour un système où toutes les particules sont différentes, est toujours nulle.

5.3 Discussion

Les discussions sur les problèmes théoriques et épistémologiques sur les différentes versions de l'entropie statistique et du formalisme qui les encadrent ont été nombreuses, se poursuivent aujourd'hui, et la plupart de ces problèmes sont liés entre eux. Peu de discussions toutefois prennent la peine de distinguer ces différentes versions (comme précédemment) ou d'énoncer ce qui est recherché par de telles définitions. L'objectif est ici de poursuivre l'examen critique entrepris précédemment et ainsi de se positionner dans les débats épistémologiques entourant le concept d'entropie dans sa version statistique. (Pour des discussions approfondies sur plusieurs enjeux discutés ici, voir principalement Ehrenfest & Ehrenfest 1912 ; Sklar 1993 ; Dugdale 1996 ; Callender 1999 ; Albert 2000 ; Uffink 2007 ; Frigg 2008 ; Frigg & Werndl 2011.)

5.3.1 Connexions

Comme on peut le présumer, ces différentes versions de l'entropie statistique peuvent être reliées entre elles²⁵. Il importe d'abord de remarquer que $S_{B,c}$ est une fonction d'une distribution d'une partition sur l'espace- μ ($\Gamma\mu$), tandis que $S_{B,m}$ utilise plutôt une partition sur l'espace- γ ($\Gamma\gamma$). Le point important est de réaliser que chacune de ces distributions correspond à une région bien définie de $\Gamma\gamma$. En effet, puisque l'espace- γ est le produit cartésien de N copies de l'espace- μ , une partition sur l'espace- μ induit aussi une partition sur l'espace- γ .

1. Ainsi, dans le cas de l'entropie de Boltzmann, si la mesure des cellules ω_i est $\delta\omega$ et donc invariante, c'est-à-dire la même pour chaque cellule, alors

$$\mu(\Gamma_M(x)) = (N/n_1!, \dots, n_m!) (\delta\omega)^N = \Omega(\delta\omega)^N,$$

de sorte que $k \ln \mu(\Gamma_M(x)) = k \ln \Omega(\delta\omega)^N = k \ln \Omega + k \ln (\delta\omega)^N$. Et puisque $k \ln (\delta\omega)^N (= kN \ln \delta\omega)$ est constant, $S_{B,m}(M)$ égale donc $S_{B,c}$ à une constante près.

2. En outre, en supposant comme précédemment que $\mu(\Gamma_M(x)) = \Omega(\delta\omega)^N$, et si $n_i \gg 1$ pour tout i , avec l'approximation de Stirling on obtient $\ln \mu(\Gamma_M(x)) \approx N \ln N - \sum n_i \ln n_i + N \ln (\delta\omega)$, et si l'on pose $n_i = N f(x, i) \delta\omega$, alors $\ln \mu(\Gamma_M(x)) \approx -kNH[f_i]$ et donc $S_{B,m}(M)$ est approximativement égale à $S_{B,f}$ pour les grands systèmes (où N est grand).

²⁵ Cette section est plus formelle et ceux ayant encore des frissons à l'idée de leurs cours de mathématiques peuvent passer à la prochaine.

3. Tel que mentionné, $S_{B,c}$ peut se ramener à $S_{G,f}$ lorsque la probabilité des micro-états est constante, soit lorsque $\rho(x) = p_i = \text{constante}$. Autrement dit, les entropies de Boltzmann et de Gibbs coïncident lorsque la distribution de probabilité ρ est une distribution uniforme des états macroscopiques auxquels appartient le micro-état x . Donc, l'entropie de Boltzmann $S_{B,c}$ peut être vue comme un cas particulier de l'entropie de Gibbs $S_{G,f}$. Mathématiquement parlant, l'entropie de Gibbs est une extension de l'entropie de Boltzmann, car elle place la définition de cette dernière au niveau des distributions de « l'espace des phases réduit » (Maes & Netocny 2003 : 275). Toutefois, le passage d'une approche traitant de systèmes individuels, à la Boltzmann, à une approche d'ensemble de systèmes, à la Gibbs, doit être justifié.

5.3.2 Réduction

L'affirmation à l'effet qu'une théorie se réduit à une autre, ou survient, ou émerge d'une autre, peut être, dit simplement, le fruit d'une recherche fort compliquée et le début d'une longue série de discussions philosophiques (Chapitre 2). Car la supériorité d'une théorie par rapport à une autre est conditionnée par les objectifs poursuivis avec une théorie, lesquels peuvent être incompatibles voire antagonistes.

Il y a au moins deux propriétés fondamentales à la version thermodynamique du concept d'entropie qui sont censées être représentées au sein de la version statistique : (i) la *non-décroissance* pour les processus adiabatiques (sans échange de chaleur) et (ii) l'*additivité* où l'entropie de deux systèmes est la somme de deux entropies individuelles. La raison du critère d'additivité vient de la définition de la température $T = \partial E / \partial S$, car si l'entropie n'avait pas la propriété d'additivité, alors le fait de coupler deux systèmes en un seul, doublant ainsi E , V et N , ferait en sorte de doubler la température, ce qui n'est évidemment pas observé. Le critère de non-décroissance est, bien entendu, dérivé du principe entropique stipulant que l'entropie d'un système isolé, soit son entropie interne, ne peut qu'augmenter. Cependant, selon Penrose (1970 : 156), il serait impossible de satisfaire (i) et (ii) en mécanique statistique. En fait, la formulation de Boltzmann, $S_{B,m}$, satisferait (i) correctement mais pas (ii), tandis que ce serait plutôt l'inverse pour celle de Gibbs, $S_{G,f}$. Il serait toutefois possible de rendre ces définitions équivalentes de façon à satisfaire correctement (i) et (ii) en s'accommodant d'une part d'approximations jugées négligeables et en considérant d'autre part des systèmes infinis dans ce qui est appelé la « limite thermodynamique », soit lorsque le nombre de molécules tend vers l'infini ($N \rightarrow \infty$).

Il suffit de considérer ces trois hypothèses, présentées par Callender (1999), pour mesurer l'ampleur du problème : (i) l'entropie statistique S est une fonction des variables dynamiques $x(t)$ d'un système

individuel ; (ii) $S(x(t)) = S(T_t x(t))$, où T_t est un opérateur temporel ; (iii) l'espace des phases Γ est borné, c'est-à-dire que le système est fermé. Or, si l'on prend en compte ces trois hypothèses, alors en raison des équations de la dynamique hamiltonienne S ne peut *pas* croître de manière monotone, donc elle ne peut pas satisfaire le critère de non-décroissance. Une approche commune est évidemment celle de Gibbs qui rejette (i) par la méthodologie des ensembles. La question est alors celle de la référence et de la pertinence (déjà discutées) d'une quantité telle que $S_{G,f}$ qui renvoie à une distribution de probabilité sur un ensemble fictif.

Du point de vue de la mécanique classique, donc, l'entropie thermodynamique ne peut être « réduite » à l'entropie statistique. Soit on accuse cet « échec » et on le nomme « émergence », une façon d'étiqueter notre ignorance, soit on applique un cadre descriptif « surnuméraire » par rapport aux équations dynamiques, ou « supra-thermodynamiques » (Ridderbos 2002), que l'on nomme le « grenage grossier » (prochaine section).

5.3.3 Irréversibilité

Il y a principalement deux objections (présentées plus tôt) à propos de la prétention de l'entropie statistique à rendre compte de l'irréversibilité des systèmes macroscopiques : celle de la réversibilité (soulevée par Loschmidt) et celle de la récurrence (soulevée par Zermelo). Elles s'appuient sur les caractéristiques de la mécanique classique afin de faire ressortir leur incompatibilité avec l'argumentation dont est issue l'entropie statistique. L'ingrédient que l'on retrouve le plus fréquemment dans les tentatives de solution à ces objections est de nature probabiliste. Cet ingrédient est matière à controverse. Malgré tout, l'accroissement d'entropie est souvent invoquée afin de rendre compte de « faits fondamentaux » comme la directionnalité du temps, de la « flèche du temps » (Chapitre 4).

L'objection de la réversibilité souligne l'IRT des lois de la mécanique classique : si, pour un intervalle de temps donné, un processus d'un état initial vers un état final est possible, alors, en renversant les vitesses des particules et pour un même intervalle de temps, un processus passant de l'état final à l'état initial est aussi possible. Un calcul « vers le futur » montre que la majorité des trajectoires des micro-états, pour un macro-état initial donné correspondant au présent, évolueront vers d'autres micro-états dans une région de l'espace des phases associée à un macro-état final, tandis qu'un calcul « vers le passé » montre que la majorité des trajectoires des micro-états, pour ce même macro-état donné au moment présent, ont évolué d'une région de l'espace des phases associée à un certain macro-état initial. Selon la présentation du chapitre précédent, l'objection stipule que, compte tenu des lois de

la mécanique classique, si un système est dans un état de non-équilibre au temps t_2 , alors il très probable que

(1) l'état de non-équilibre au temps t_2 évoluera vers un état d'équilibre au temps t_3 ,

(2) l'état de non-équilibre au temps t_2 a évolué d'un état d'équilibre au temps t_1 ,

où $t_1 < t_2 < t_3$. Malgré tout, des phénomènes tels que (2) ne sont jamais observés. Ainsi, l'objectif initial d'expliquer les phénomènes observés ne semble pas être rencontré.

Suivant le postulat statistique reliant volume en phase et probabilité, cela signifie que l'entropie de tout système thermodynamique dans un quelconque macro-état est aussi susceptible d'augmenter vers le futur que vers le passé, c'est-à-dire que l'entropie était en train de diminuer jusqu'au moment présent. Autrement dit, si un système évolue d'un macro-état associé à un faible volume en phase, donc de faible entropie, vers un macro-état associé à un grand volume en phase, donc d'entropie élevée, une rétrodition basée uniquement sur ce dernier macro-état et sur la mécanique classique conclura que l'entropie passée était encore plus élevée. Il est donc possible *en théorie* (pas nécessairement en pratique) d'obtenir un processus microscopique qui est impossible au niveau macroscopique alors que le premier est censé rendre compte du second. Tel que mentionné, le problème n'est pas tant que cette approche décrive des phénomènes qui ne se produisent pas, mais plutôt qu'elle fasse des prédictions qui soient falsifiées.

L'objection de la récurrence est sans doute plus incisive pour la stratégie explicative de l'entropie statistique. Comme vu précédemment, elle repose sur le théorème de Liouville d'une part, et sur celui de Poincaré, d'autre part, qui en est une conséquence. Le théorème de Liouville stipule en substance qu'un volume de l'espace des phases soumis aux lois de la mécanique classique demeure constant ou invariant dans le temps, même s'il se déforme. Or, le théorème de Poincaré pose une région composée de *toutes* les autres régions de l'espace des phases disponible pour un système dont l'énergie est fixe, il s'ensuit selon le théorème précédent que le volume de cette très grande région demeure le même ; donc la région de l'espace des phases initiale est la même que la région finale après une certaine évolution temporelle. Par conséquent, un système dont l'énergie est fixe et défini par une région finie de l'espace des phases doit retourner exactement à son état initial et ce, avec une récurrence arbitraire (excepté pour un ensemble d'états singuliers de mesure zéro).

Le problème soulevé par ces objections, lequel prend la forme d'un dilemme cornélien, est donc le suivant : soit l'entropie statistique est soumise aux contraintes de la mécanique classique, auquel cas elle ne peut présenter un caractère d'irréversibilité défini par une croissance dans le temps, c'est-à-dire vers le futur et donc une asymétrie temporelle telle que l'exprime le second principe, soit elle manifeste

vraiment un tel caractère d'irréversibilité et ne peut donc pas être compatible avec les équations de la mécanique classique. De fait, avec les systèmes déterministes soumis au théorème de Liouville, il a été vu que la distribution de probabilité ρ est invariante dans le temps ($d\rho/dt = 0$) et que, conséquemment, il est impossible pour une distribution d'évoluer vers une distribution caractérisant un état d'équilibre, soit la distribution d'ensemble microcanonique. Ce point est important car il demeure possible de postuler un principe (d'évolution probable) stipulant que tout état est plus probable que l'état qui le précède, et définir l'entropie en fonction d'une certaine mesure de probabilité, en sorte que la valeur d'entropie d'un état soit toujours plus élevée que celle de l'état qui le précède. (En fait, il s'agit d'une « trivialisation » de l'augmentation de l'entropie (section 6.3.2.1).) Mais aucune description de l'évolution de l'entropie dans le temps n'est fournie.

Une première tactique pour éviter le problème (au sens « de se dérober » ou de « changer de direction ») est d'accepter les termes apparemment contradictoires du problème, c'est-à-dire de tenter, sans égard à la possibilité même de le faire, de dissoudre ce lieu d'antinomies où les thèses en présence s'affrontent de manière irréductible. Ainsi, on peut bien prendre acte de cette irréductibilité et afficher un agnosticisme se dérochant à toute conception modale du problème, donc à toute explication du phénomène. C'est-à-dire, une telle approche se limite à la description de différents « niveaux », à défaut de mieux pour l'instant, comme celui des particules régies par les équations réversibles de la mécanique classique et celui de la thermodynamique faisant état d'une irréversibilité. Sans aller plus loin. En effet, une antinomie ne « prouve » rien d'elle-même (Sorensen 2009 : 110). Aller plus loin et proposer une explication serait se livrer à une illusion transcendantale selon l'expression kantienne. Peu importe, tel que mentionné, d'aucuns ont proposé que l'entropie, et plus précisément sa croissance exemplifiant l'irréversibilité, serait une *propriété émergente*. Mais l'antinomie demeure. Et le caractère émergent est inféré justement de l'irréductibilité des thèses du problème apparemment antinomique. En ce sens particulier, l'émergence est un concept « négatif » découlant de l'échec d'une explication ou d'une réduction (Chapitre 2).

Une autre tactique générale pour éviter ces problèmes est d'avoir recours à la *méthodologie du grenage grossier*. Il s'agit d'établir une partition de l'espace des phases en petites boîtes ou cellules de dimensions fixes. Plus précisément, l'idée est de faire la moyenne sur l'espace des phases pour chaque cellule de dimensions finies par une application mathématique²⁶. Le grenage grossier se présente sous deux formes, l'une par rapport à l'espace- μ chez Boltzmann (présenté précédemment) et l'autre par rapport

²⁶ $GG : \rho(x) \rightarrow GG \rho(x) = \underline{\rho}(\omega, p, q, t) = \sum \rho^*(i) 1_{\omega_i}(x)$, où $\rho^*(i) = \int \omega_i \rho(x) dx / \int \omega_i dx$, et 1 représente la fonction caractéristique : $1_A(x) = \{1 \text{ si } x \in A ; 0 \text{ ailleurs, où } A \text{ est une région de l'espace des phases. (Pour plus de détails, voir Ridderbos 2002 ; Uffink 2007.)}$

à l'espace- γ chez Gibbs. Pour ces deux formes l'idée générale est que, alors que le système évolue dans le temps, son micro-état évolue vers une distribution uniforme dans un grand nombre de boîte ou encore que son volume se « répand » dans plusieurs d'entre elles. En conséquence, l'entropie augmente et l'irréversibilité recherchée serait recouvrée. (Pour plus de détails, voir Sklar 1993 ; Schulman & Gaveau 2001 ; Ridderbos 2002 ; Frigg 2008a.)

Mais comment justifier une telle procédure ? Dans un premier temps, la partition de l'espace des phases ainsi obtenue n'est *pas* soumise au théorème de Liouville et par conséquent elle n'est pas soumise à l'invariance temporelle. Par contre, l'objection de la réversibilité tient toujours car si l'on suit l'évolution inverse du système alors on constate une augmentation vers le passé de l'entropie définie par le grenage grossier. Ainsi, bien qu'elle réponde au problème de l'invariance temporelle, elle n'offre pas de moyens de rétrodiction adéquats. Autrement dit, ce que décrit la théorie en cet état est en contradiction avec ce qui est observé. On pourrait alors remettre en cause la pertinence d'une telle procédure. Le principal argument pour son application repose sur l'idée qu'une mesure se limite à ce qui est observable et qu'en conséquence il est impossible de déterminer le micro-état avec une précision absolue. Ainsi, selon cet argument, cette limitation dans ce qui nous est possible de connaître par l'observation devrait être prise en compte dans la description des systèmes physiques à l'étude. L'objection évidente à ce type d'argument renvoie à la subjectivité, que ce soit par les limites de résolution des appareils de mesure ou les limites cognitives de l'observateur, que cela introduit dans la description du système physique.

L'objection quant à la subjectivité n'est toutefois pas entièrement légitime (Goldstein 2001 ; Ridderbos 2002). D'abord, il est davantage question ici d'un pluralisme de partitions possibles dont l'indétermination présente un aspect arbitraire et non pas d'états mentaux ou de préférences d'un quelconque observateur. Il n'est pas question de relativiser la description du système par rapport à l'état de connaissance de cet éventuel observateur. Une fois que la partition est déterminée, ou si l'on préfère, sélectionnée, l'évolution temporelle de la distribution ainsi établie est indépendante de la résolution de mesure de l'observateur. Le problème est donc tout à fait semblable à celui rencontré dans la discussion sur les probabilités (section 5.1.3). Il n'y a pas, à la base, de règles formelles claires et bien établies afin de procéder à l'attribution de valeur de probabilité ou d'une mesure de probabilité à une gamme de possibles. À moins, bien sûr, de faire appel à des principes comme celui de la raison insuffisante, menant à l'équiprobabilité, ou encore celui identifiant la fréquence relative à la probabilité. Mais ces principes ne sont pas suffisamment solides pour encadrer l'appareil probabiliste. Malgré cet aspect arbitraire, on l'a vu, cela ne suffit pas à justifier l'interprétation subjectiviste.

Dans un second temps, cette procédure, face à cette indétermination, doit justifier son choix de la « bonne » partition. En effet, la mécanique classique n'offre aucun expédient sur la taille des cellules ni sur la nature de leurs dimensions. Par exemple, il n'y a rien qui permette de justifier une préférence pour une partition « position-moment », comme chez Boltzmann, plutôt qu'une partition sur l'énergie, par exemple. Comme il y a plusieurs partitions possibles, l'augmentation d'entropie n'est pas unique. Certains paramètres intensifs comme la température recevraient alors plus d'une définition puisqu'ils sont définis par rapport au taux de variation de l'entropie. Le problème de la taille des cellules peut être résolu si l'on autorise l'apport de la mécanique quantique qui définit la précision maximale dans l'espace position-moment avec le principe d'indétermination d'Heisenberg. Ce principe permet aussi de statuer sur une impossibilité théorique de spécification absolue d'un micro-état ou microcondition. Un tel apport n'est toutefois pas justifié, tel que mentionné, dans le projet initial de la mécanique statistique. (Pour une discussion de l'apport de la mécanique quantique en mécanique statistique, voir Tolman 1938 ; Zeh 1989 ; Gallavotti 1999 ; Balian 2005 ; Yoshioka 2007 ; Tuckerman 2010.)

La sous-détermination d'une partition particulière est un problème dans le cadre d'une explication de la thermodynamique par la mécanique statistique. Plus précisément, la question est de savoir s'il est possible ou non d'expliquer les phénomènes thermodynamiques observés par la théorie de la mécanique classique et quelques postulats probabilistes jugés raisonnables. Que la mécanique statistique soit incapable de justifier le choix d'une partition en particulier n'implique pas que cette partition n'existe pas. De plus, différents choix de partition mène souvent aux mêmes prédictions observables. Mais, comme dans le cas, déjà discuté, des probabilités, cette partition est généralement considérée comme arbitraire ou subjective, ou à tout le moins anthropocentrique. Car elle suppose qu'aucun observateur n'a d'accès empirique aux valeurs des micro-états, et c'est le cas de la mécanique statistique en général. Cette inaccessibilité empirique deviendrait ainsi une inaccessibilité, une incomplétude épistémique, une cause de certaines limites de nos moyens d'observation, d'un « manque d'information » sur l'état microscopique, ce qu'est censé illustrer le « démon de Maxwell » (section 5.3.4.3). Comme la définition du concept statistique de l'entropie exige, en quelque sorte, une incomplétude épistémique, il s'ensuit, selon cette ligne de pensée, que ce concept est relatif à cette incomplétude épistémique, à ces connaissances incomplètes, et (donc) au sujet.

Bien que la procédure du grenage grossier semble, en partie du moins, arbitraire, au moins autant que peut l'être l'inférence inductive s'appuyant sur un nombre *prétendument* suffisant pour dire que telle fréquence relative *est* une valeur de probabilité, elle n'est pas étrangère au « découpage », pour reprendre le même terme, de nos expériences qui sont par définition macroscopiques. C'est-à-dire que lorsque nous observons un système changer, comme un corps passant d'un état de non-équilibre à un état d'équilibre, évidemment ses propriétés observables changent, et il est raisonnable de dire que son

micro-état change aussi ; mais il paraît aussi raisonnable de dire que l'inverse n'est pas vrai, c'est-à-dire qu'il semble possible que son micro-état change sans que ses propriétés observables changent. Ce n'est qu'une conséquence, qu'une autre manière de dire que le micro-état est inobservable. Car si tel est le cas, le micro-état peut changer sans qu'il n'y ait de changement observable. Après tout, on ne suppose pas que tout phénomène soit accessible à nos modestes moyens d'observation. Néanmoins, les paramètres macroscopiques et donc observables définissent partiellement, en ce qu'ils offrent un cadre formel, cette partition, ce « découpage » de l'espace des phases.

L'irréversibilité thermodynamique n'est donc pas une *conséquence* de nos limitations cognitives ou observationnelles, et donc du sujet observateur. Une tasse de café va se refroidir peu importe qui l'observe. En revanche, la description de l'évolution temporelle de l'entropie statistique exige une partition en termes observationnels qui rendent compte de la phénoménologie thermodynamique, laquelle est relative, sans en dépendre totalement, à des capacités d'observation, mais pas à de quelconques états mentaux. L'aspect subjectif est donc partiel. Peut-être que cette description en termes observationnels rend compte du même coup des possibilités physiques du système au niveau macroscopique comme la capacité à effectuer un travail, elles-mêmes exprimées par la phénoménologie thermodynamique. En outre, il a déjà été vu (Chapitre 4) que l'irréversibilité thermodynamique, au sens d'irrécupérabilité, n'était pas absolue au sens où une intervention extérieure pouvait permettre de retrouver l'état initial d'un certain système. Pour que cette intervention soit efficiente en ce sens, il faut qu'elle soit coordonnée à une quantité d'information suffisamment grande pour tenir compte de tout le sous-système, de tous les composants, toutes les particules du système macroscopique. Un système macroscopique comme une tasse de café, donc, pourrait ne pas se refroidir *si* l'information sur son sous-système est suffisante et *si* une intervention sur celui-ci est adéquate. Dans ce cas, il s'agit d'empêcher le système de se refroidir et donc l'évolution vers l'état d'équilibre n'est pas une fatalité.

5.3.4 Interprétations

Face à un concept aussi « prodigieusement abstrait », selon les mots de Poincaré, il n'est pas étonnant que les interprétations se soient multipliées. À titre d'exemple, Sethna (2011 : 77) affirme dans le même paragraphe que l'entropie est une mesure du « désordre », de notre « ignorance à propos d'un système » ainsi que des « changements irréversibles se produisant dans un système ». Autre exemple, selon Prigogine & Stengers (1979 : 195), en identifiant l'entropie au nombre de complexions, Boltzmann a interprété « la croissance irréversible de l'entropie comme l'expression de la croissance du désordre moléculaire, de l'oubli progressif de toute dissymétrie initiale car toute dissymétrie est

improbable par rapport à l'état correspondant au nombre de complexions maximal ». (Voir la section 5.1.2 pour quelques considérations générales sur les interprétations.)

5.3.4.1 PROBABILITÉ

L'interprétation probabiliste découle directement du projet même de la mécanique statistique. Il a été vu que ce qui permet de caractériser un système macroscopique, son macro-état, l'« information » disponible sur ce dernier, est compatible avec une multitude de scénarios caractérisant ce même système, mais dans un autre « langage », avec d'autres « informations », celui de la mécanique²⁷. Puisque le système est considéré avoir un seul état peu importe le « langage » utilisé pour le décrire, que ce soit ici la thermodynamique ou la mécanique, il s'ensuit que cette multitude est décrite comme une gamme des possibles où le système se trouve dans l'un ou l'autre des scénarios possibles, dans l'un ou l'autre des micro-états possibles, parce que compatibles avec les données, les informations disponibles. Ces données ou ces informations, bref ce qui détermine les valeurs des paramètres thermodynamiques et macroscopiques, déterminent aussi le cadre formel dans lequel s'inscrit cette gamme des possibles, au demeurant inobservables et microscopiques. Un état macroscopique peut donc être dans l'un ou l'autre de ces micro-états, ou encore il peut être dans telle ou telle région de l'espace des phases. Il semble donc naturel de définir ce rapport entre ce qui détermine l'état macroscopique et ce qui détermine la gamme des possibles microscopiques comme la probabilité de cet état macroscopique. De plus, compte tenu du *fait* que les systèmes thermodynamiques tendent à évoluer vers un état d'équilibre *et* qu'un tel état est typiquement décrit comme un état compatible avec la plus grande gamme de possibles, le plus grand nombre de micro-états possibles ou la plus grande région de l'espace des phases, et de loin, alors il semble aussi naturel d'expliquer cette tendance fortement corroborée comme un processus d'un état improbable à un état plus probable.

Cette gamme des possibles peut aussi être décrite par des ensembles plus ou moins grands circonscrivant ces possibilités. Dans l'exemple du tirage d'une pièce de monnaie, la gamme des possibles est généralement définie par deux ensembles contenant chacun une possibilité, en l'occurrence « pile » ou « face », mais elle peut aussi être élargie en spécifiant, par exemple, la position de la pièce sur la table. Contrairement à l'exemple de la pièce de monnaie toutefois, le micro-état en mécanique statistique n'est pas en principe, ni en pratique donc, accessible.

La quantité Ω est associée à une équiprobabilité, où chaque micro-état a le même « poids » par rapport à la quantité totale de micro-états possibles, $\mu(\Gamma_M)$ mesure le volume en phase d'une région de

²⁷ Certes, cette description est très ébauchée pour l'instant, mais ne sert qu'à introduire les prochaines discussions, plus étoffées, sur les interprétations du concept d'entropie.

l'espace des phases, donc définit une gamme de possibles, tandis que ρ est une distribution de probabilité. Il n'est donc pas surprenant, encore une fois, qu'on ait interprété l'entropie comme une mesure de probabilité. Mais ce qu'il faut ici est une justification de cette connexion, de cette interprétation de l'entropie en termes probabilistes. Il a été vu précédemment (section 5.1.3) que l'attribution d'une valeur de probabilité à un système ou un phénomène n'était pas encadrée par des règles formelles comme peut l'être le calcul des probabilités. Malgré la possibilité d'établir une gamme des possibles par des mesures empiriques ou par un traitement nomologique, une mesure de probabilité, établissant une valeur de probabilité dans un cas particulier, n'est pas pour autant définie. Il serait possible de le faire, par exemple, s'il était justifié de procéder à l'identification de cette valeur pour un phénomène donné avec sa fréquence relative. Ou encore si le principe d'indifférence était complètement justifié²⁸. Bien entendu, ces deux procédures sont couramment utilisées et il ne fait aucun doute que leur emploi soit parfois justifié. Mais ce n'est pas toujours le cas et la définition de leur contexte de justification est (encore) lacunaire. Ainsi, il est possible de connaître, par une mesure par exemple, l'état macroscopique d'un système sans que l'on soit pour autant justifié d'établir une valeur de probabilité à son micro-état.

Davey (2008) affirme ainsi que tout principe de la forme « si l'on connaît l'état macroscopique d'un système alors on est justifié de décrire son micro-état avec telle mesure de probabilité » est nécessairement faux. Certes, on le répète, il a été soutenu plus tôt (section 5.1.3) que le choix d'une mesure de probabilité ne faisait pas l'objet d'un critère formel. Mais en rejetant ce genre de « principe », c'est-à-dire en exigeant un tel niveau de justification, c'est *toute* attribution de valeur de probabilité ou de mesure de probabilité qui échappe au domaine de ce qui peut être « justifié ». Ce n'est pas seulement la connexion entre probabilités et entropie qui est en péril, mais celle entre probabilités et tout système ou processus physique. Dans son article, il soutient qu'il n'y a pas de principe simple, apriorique, permettant d'inférer avec une justification suffisante la probabilité d'un macro-état sur la simple base de l'observation établissant les paramètres de macro-état et les équations de la mécanique classique. Son argument visant à briser la connexion entre entropie et probabilité s'appuie sur une expérience de pensée (une autre) où, par décret gouvernemental, les citoyens doivent, dans un cas, tout faire pour garder les verres d'eau à moitié plein de glace, et dans l'autre, ils doivent tout faire pour garder les verres d'eau exempts de glace. Dans le premier cas, donc, lorsque la glace d'un verre se met à fondre, le système qu'est le verre passe d'un état probable à un état moins probable avec une entropie élevée, tandis que dans le second cas, le verre passe d'un état peu probable à un état

²⁸ Halpern (2006) tente de justifier l'équiprobabilité par la « méthode des mondes aléatoires » (« *random-worlds method* ») consistant à offrir une sémantique aux degrés de croyance selon une distribution de probabilité sur un ensemble de mondes possibles. Cette méthode ne semble pas vraiment différente toutefois du principe d'indifférence.

plus probable avec, aussi, une entropie élevée. Ces contingences (le décret gouvernemental) empêcheraient une conclusion univoque entre un état d'entropie élevée et une haute probabilité.

Je souhaite présenter trois arguments montrant non pas que la connexion entre entropie et probabilité est nécessairement vraie, mais plutôt que l'argument de Davey est insuffisant pour soutenir la thèse de l'impossibilité d'une conclusion univoque entre un état d'entropie élevée et une haute probabilité. Le premier de ces arguments concerne l'omission, dans son expérience de pensée, de tenir compte de l'isolation du système, c'est-à-dire des verres d'eau. Si, en effet, un système comme un verre d'eau n'est pas isolé et qu'il peut échanger matière et énergie avec son environnement, alors son entropie peut augmenter ou diminuer. C'est précisément cette condition qui permet d'évaluer correctement le comportement de systèmes ouverts comme un réfrigérateur ou un être vivant. Or, comme les verres d'eau dans le précédent exemple reçoivent matière et énergie, puisqu'on leur donne ou enlève de la glace, on ne peut conclure de manière univoque à un changement d'entropie en faisant abstraction de ces échanges.

L'importance de définir le système adéquatement peut être illustrée de manière plus précise à l'aide d'exemples vus plus tôt (Chapitre 4). En effet, si le système n'est pas bien défini les inférences quant à ce qui possible ou non pour ce système risquent fort d'être invalides. Ainsi, dans les discussions à propos de l'irréversibilité des processus thermodynamiques, Uffink (2001) soutient que ni la théorie de Carnot ni celle de Planck ne stipulent l'irréversibilité au sens d'irrécupérabilité des conditions initiales. Un système qui échange de la chaleur avec un autre système voit son état changer et il ne reviendra pas à son état initial *de lui-même*. Il est cependant possible d'intervenir et de lui fournir de la chaleur, par exemple, en sorte qu'il revienne effectivement à son état initial. C'est pourquoi l'irrécupérabilité de ce processus peut être définie si l'on y pose une condition de principe, soit l'impossibilité de recouvrer l'état initial du système *à partir du système*. L'intervention précédente, provenant de l'environnement, serait donc exclue et ce processus pourrait être qualifiée d'irréversible. Un autre exemple, plus spécifique, vient de l'application de la thermodynamique aux problèmes environnementaux actuels. L'énergie est partout et par dessus tout elle se conserve. Ce n'est donc pas une crise de l'énergie mais une « crise de l'entropie » (Deutscher 2008). C'est-à-dire que la demande mondiale concerne une énergie sous forme de travail, de faible entropie ou d'exergie (Wall & Gong 2001), et compte tenue de l'augmentation d'entropie lors des transformations d'énergie, un retour à l'état initial de faible entropie exige d'ouvrir les limites du systèmes, donc de consommer davantage d'énergie et de ressources. En somme, si les limites du système ne sont pas spécifiées, on peut difficilement établir ce qui est possible ou non pour ce système, ni le qualifier d'irréversible ou non.

Le second argument a trait à la bonne définition de l'espace des phases, qui est nécessaire pour déterminer correctement l'entropie statistique. Tel que mentionné, si les frontières définissant le système ne sont pas claires, parce qu'il serait en interaction avec l'environnement, alors la valeur d'entropie en est affectée. Cette question est donc corrélative à la question précédente de la bonne définition du système. Conséquemment, la valeur de probabilité attribuée en fonction du nombre de complexions reste indéfinie ou erronée lorsque le système est loin de l'équilibre, dans le cas de systèmes ouverts. Ainsi, les phénomènes d'auto-organisation, qui sont pourtant légion, se voient attribués une probabilité nulle. Dans pareils cas, « l'application du concept de probabilité fondé sur le calcul du nombre de complexions devient [...] impossible » (Stengers & Prigogine 1979 : 215).

Le troisième argument concerne l'attribution d'une probabilité aux macro-états des verres d'eau. Pour qu'une conclusion valide puisse être tirée quant à la soi-disant connexion entre entropie et probabilité il faut, trivialement, que ces deux concepts soient appliqués aux mêmes systèmes. Or, dans l'exemple des verres d'eau, on parle de l'entropie de systèmes individuels d'une part et de probabilité d'une classe de référence d'autre part. C'est-à-dire que la probabilité que l'on attribue à un macro-état est déterminée à partir de son appartenance à une classe de référence de systèmes *similaires*. La gamme des possibles établissant la probabilité est donc déterminée selon cette similitude, en l'occurrence un verre d'eau dans un monde soumis au décret gouvernemental. Mais cette similitude n'est pas adéquate ici car les conditions initiales biaisent les résultats, elles introduisent des interactions causales affectant la gamme des possibles, donc la probabilité. Par exemple, les résultats définissant la gamme des possibles ainsi que la fréquence relative du tir d'un dé sont affectés par l'impulsion initiale, laquelle doit être aléatoire, comme quoi des tirs verticaux sur la face « 1 », par exemple, risquent justement d'avoir une influence induite. Aussi, je peux bien calculer la probabilité d'avoir contracté un virus étant donné que j'ai une fièvre de 39°C, laquelle ne sera pas nécessairement la même que celle d'avoir contracté un virus étant donné que je fréquente tel hôpital ou que je suis membre de tel club, etc. Ainsi, ces probabilités conditionnelles ne seront pas les mêmes. Dans l'exemple des verres d'eau, l'interaction entre ces verres et les bonnes gens respectant le décret gouvernemental n'est pas anodine. Évidemment, la probabilité qu'un verre soit rempli de glace n'est pas la même si ce verre est en interaction avec son environnement duquel il reçoit de la glace, s'il se trouve en Arctique, etc.

Mais l'argument de Davey demande un peu plus d'acuité. Car il ne s'agit pas seulement de souligner, trivialement, que différentes classes de référence mènent à différentes valeurs de probabilités. Car son argument stipule que certaines contingences, comme le décret gouvernemental, impliquent qu'une seule classe de références circonscrivant les macro-états de certains systèmes, comme les verres d'eau, soit raisonnable. Le problème n'est donc pas d'identifier la bonne classe de référence, celle où les membres seraient suffisamment similaires. La probabilité ici n'est que la fréquence relative d'une

certaine propriété ou d'un ensemble de propriétés définissant le macro-état faisant l'objet de cette probabilité, soit le nombre de verres d'eau avec (ou sans) glace parmi l'ensemble des verres d'eau. Le problème est ailleurs. L'argument échoue car, même si la probabilité des macro-états peut être bien définie en supposant que la fréquence relative joue bien le rôle qu'on veut lui faire jouer, l'entropie statistique associée à ces macro-états ne l'est pas et par conséquent on ne peut inférer que la connexion entre probabilité et entropie est brisée. En effet, comme les systèmes, les verres d'eau, ne sont pas adiabatiquement isolés, leur hamiltonien n'est pas indépendant du temps ou, ce qui revient au même dans ce cas, leur espace des phases n'est pas défini entièrement par les strictes variables canoniques des molécules des verres. Une objection visant à maintenir la validité de l'argument consisterait à dire que les verres d'eau sont isolés entre deux interactions avec ces bonnes gens obéissant au décret gouvernemental. Mais cette objection n'évite pas notre troisième argument, car, justement, si on fait abstraction de ces interactions, alors on constatera que la vaste majorité des verres d'eau tendent effectivement vers une entropie plus élevée.

La principale raison pourquoi Davey (2008) se permet de faire abstraction des interactions entre les systèmes et leur environnement découle de l'hypothèse selon laquelle l'entropie ne dépend que de son macro-état (voir aussi Davey 2011). Mais ces interactions font évidemment en sorte que le système n'est pas isolé et par conséquent la définition de son espace des phases doit en tenir compte. Alors comment établir cette connexion entre probabilité et entropie ? Soit trois hypothèses empiriques : (i) les états de faible entropie sont rarement observés ; (ii) les états de faible entropie ne sont jamais observés à partir de l'évolution de certaines conditions initiales ; (iii) lorsqu'un état de faible entropie est effectivement observé, on peut toujours trouver ou inférer une certaine interaction entre ce système et son environnement, c'est-à-dire qu'il n'est pas ou n'était pas isolé dans un proche passé. Si, nonobstant (i), un système dans un état de faible entropie est observé, alors on peut bien inférer (iii). Toutefois, si l'on conteste (iii), il faut alors supposer que ce système est depuis longtemps dans cet état, qu'il se trouve en quelque sorte dans un état d'équilibre *et* de faible entropie. Ce qui revient à contester (ii). Mais (ii) ne pourra pas être falsifiée, si tant est que l'on puisse se fier aux observations précédentes, que l'on puisse inférer, *jusqu'à preuve du contraire*, de ses multiples corroborations sa non-falsification future, pour ainsi dire. Accepter (ii) est tout à fait compatible avec la mécanique classique qui stipule que la vaste majorité des conditions initiales, des micro-états, mène à une entropie plus élevée au sens qu'en donnent Boltzmann et Gibbs.

Mais parler de la probabilité d'un macro-état, par exemple de la probabilité que la glace fonde, est inapproprié, « captieux » (Leeds 2003 : 128). La discussion précédente montre plutôt que ce qui est très probable est que le système se trouve dans un micro-état, dans des conditions initiales dont l'évolution déterministe selon les équations dynamiques garantit que l'entropie va augmenter, ou, en l'occurrence,

que la glace va fondre. Ainsi, dire que le système passe d'un état peu probable à un état plus probable n'a de sens que si l'on se réfère à la probabilité synchronique *abstraction faite* du processus reliant ces deux états. Comme le fait remarquer Planck (in Kuhn 1978 : 54) :

Le calcul des probabilités peut servir, si rien n'est connu d'avance, à déterminer l'état le plus probable. Mais il ne peut servir, si un état [initial] improbable est donné, d'évaluer le prochain [état]. Cela est déterminé non pas les probabilités mais par la mécanique. Soutenir que le changement dans la nature procède toujours [d'états] peu probable vers des états plus probables serait complètement sans fondement.²⁹

Puisque l'évolution est déterministe, il ne fait aucun sens de dire que le système a une « tendance causale », une probabilité d'évoluer vers un état de plus haute entropie. Étant déterministe, chaque micro-état spécifique, chaque microcondition initiale évolue inmanquablement vers le micro-état ou microcondition final déterminé par les équations dynamiques. C'est pourquoi la mécanique statistique est incompatible avec l'interprétation propensionniste, où un état initial est compatible, étant diachroniquement sous-déterminé, avec plusieurs états finaux (Clark 2001).

Ainsi, un système thermodynamique isolé a une forte probabilité synchronique (telle que définie précédemment) d'être dans un micro-état évoluant de manière déterministe vers un micro-état associé à une entropie statistique élevée. La caractérisation de ces micro-états est une autre histoire. Dire, par exemple, que les « dissymétries initiales » sont improbables ou encore que l'état le plus probable accessible à un système est celui où les événements en foule qui se produisent simultanément dans ce système *compensent statistiquement* leurs effets » (Prigogine & Stengers 1979 : 196, italiques originaux) demanderait une discussion encore plus approfondie. Néanmoins, en projetant l'image d'une partition de l'espace des phases où chaque cellule peut être occupée par une ou plusieurs particules et en stipulant que toute chose est égale par ailleurs, ce qui revient d'une certaine façon à endosser l'équiprobabilité, on peut imaginer cette dissymétrie correspondre à un petit nombre de cellules occupées et lui attribuer une faible probabilité.

Ces précisions faites, cette probabilité synchronique est définie à partir de la définition du système ayant un espace des phases déterminé, et ce n'est pas le cas dans l'expérience de pensée de Davey (2008). Cependant, je le répète, la définition de cette probabilité, comme toute probabilité appliquée aux phénomènes naturels, et cela Davey le note à juste titre, ne peut découler d'une mesure de probabilité parfaitement justifiée, au sens où ni la fréquence relative, ni le principe d'indifférence ne

²⁹ *Probability calculus can serve, if nothing is known in advance, to determine the most probable state. But it cannot serve, if an improbable [initial] state is given, to compute the following [state]. That is determined not by probability but by mechanics. To maintain that change in nature always proceeds from [states of] lower to higher probability would be totally without foundation.*

sont exempts d'arbitraire en l'absence de règles claires les encadrant. En cela l'assignation d'une mesure de probabilité à des systèmes ou des processus physiques est aussi problématique, aussi arbitraire que l'induction empirique. Les systèmes auxquels on attribue une valeur d'entropie ne font pas exception et c'est pourquoi ce concept n'est « subjectif » qu'en un sens très restreint.

5.3.4.2 DÉSORDRE

L'interprétation de l'entropie comme étant une « mesure du désordre » compte plusieurs adeptes (Carnap 1977 ; Wherl 1978 et 1991 ; Chandler 1987 ; Ellis 2006 ; Niven 2007 ; Müller & Weiss 2008 ; Sethna 2011). Elle trouve son origine dans les travaux de Boltzmann, qui, on l'a vu, interprète l'entropie statistique comme une probabilité à partir du nombre d'arrangements compatibles avec une distribution donnée. Mais celui-ci identifie aussi les états ordonnés à des états improbables et les états désordonnés à des états probables. D'aucuns parlent aussi de « mélangeté » (« *mixed-upness* »), d'un « caractère aléatoire » (« *randomness* ») ou d'une « quantité de chaos » (« *amount of chaos* »).

Cette interprétation n'est toutefois pas entièrement étrangère à l'entropie thermodynamique. En effet, considérant l'interprétation inspirée du second principe de la thermodynamique à l'effet que l'entropie soit une mesure de la capacité d'un système à effectuer un travail (section 5.3.4.4), et compte tenu de l'interprétation anthropomorphique d'un travail comme ce qui permet l'« instauration d'un ordre », il n'est pas absurde d'établir une correspondance entre une entropie élevée et le désordre. Cette interprétation d'« instauration d'un ordre » a deux sources. L'une d'elle est empirique, en ceci que l'intuition nous dit qu'une quantité de travail est nécessaire pour construire, bâtir, organiser ; il ne nous viendrait pas à l'idée, par exemple, de chauffer des briques pour construire une maison. La seconde est plus théorique, car elle découle de la définition microscopique d'un travail comme un transfert d'énergie « de manière cohérente » (Gould & Tobochnik 2010 : 18) ou impliquant « peu de degrés de liberté observables » (Penrose 1970 : 158).

La quantité Ω dans l'entropie combinatoire de Boltzmann ($S_{B,c}$) peut être interprétée comme une mesure de l'« ordre » d'un système composé de plusieurs éléments de manière analogue aux lettres de l'alphabet pouvant être arrangées de multiples façons différentes, et indifférentes, dans le « désordre », mais seulement d'une façon selon l'ordre habituel³⁰. Le même raisonnement vaut donc aussi pour l'entropie macroscopique de Boltzmann ($S_{B,m}$) puisqu'elle est égale, on l'a vu, à l'entropie combinatoire à une constante près si la mesure du volume des cellules en phase est invariante. Si l'ensemble des molécules présentait une très grande symétrie, si elles étaient toutes à la même vitesse, se déplaçant

³⁰ Plus précisément, de $N! = 26!$ façons *versus* la façon habituelle ou conventionnelle, soit {A, B, C, ..., X, Y, Z}.

parallèlement, et dans un petit volume, alors l'entropie serait extrêmement faible. L'association entre une configuration en phases et le concept d'ordre est donc invitante. Il importe alors de souligner le fait que la notion de désordre doit s'appliquer ici à l'espace des phases et non à l'espace tout court. Par exemple, le phénomène de cristallisation mène indubitablement à un plus grand ordre spatial bien qu'il implique aussi (ce phénomène n'étant pas entièrement compris cependant) une augmentation d'entropie. Il faut aussi que la définition d'entropie soit correctement appliquée à un espace des phases donné. Si, par exemple, on calcule l'entropie d'un jeu de carte en attribuant une probabilité à chacune des cartes, comme si les unités de l'espace des phases étaient les figures du jeu de cartes, alors il est évident que la covariation entre désordre et entropie ne tiendra pas ; car si l'on chauffe ce jeu de carte, alors l'entropie thermodynamique augmente tandis que l'ordre des cartes demeure évidemment le même.

Mais la notion de désordre n'est pas non plus une notion facile à définir ni à interpréter. On lui associe aisément des termes comme « fouillis », « confusion », « désorganisation ». Ainsi la notion de désordre n'est pas très éloignée de celle de symétrie. Par exemple, une foule lors d'un spectacle est généralement considérée comme étant dans le désordre, tandis qu'un défilé militaire est considéré comme ordonné précisément en raison des symétries qu'il présente. De manière similaire, un arrangement ou microt-état présentant des symétries dans l'espace des phases correspondra typiquement à une quantité Ω plus faible et donc à une entropie ($S_{B,c}$) plus faible.

Pour autant, la caractérisation du désordre en fonction de la notion de symétrie ne semble pas offrir une précision satisfaisante et on peut bien arguer que le concept de désordre est « vague » (Styer 2000 ; Leff 2007), « subjectif » (Denbigh 1989a ; Ben-Naim 2008 ; Michaelides 2008) ou « anthropomorphique » (Burgers 1954). Cette critique de la « subjectivité » peut être considérée comme une critique quant au caractère arbitraire de la notion de désordre. Arbitraire, d'une part, car elle n'est mesurée que par rapport à des impressions ou à des préférences personnelles sans autre orientation. C'est exactement ce que fait Ben-Naim (2008 : 10), suivant une méthode justement « subjective », qui affirme qu'il ne voit pas, « personnellement », de différence entre deux arrangements de N molécules, de volumes différents, et présentant pourtant des valeurs d'entropie différentes. Évidemment, si le concept de désordre est défini opératoirement à partir d'impressions ou des capacités d'un observateur à distinguer une forme, il est certain qu'une éventuelle critique a beau jeu de le taxer de subjectivité. Arbitraire aussi, d'autre part, puisque la mesure du désordre exige une règle qui plus souvent qu'autrement n'est établie que par convention. Denbigh (1989a : 328) donne un exemple de ce caractère qu'il qualifie aussi de « subjectif », avec l'ordonnement conventionnel et donc arbitraire dans une certaine mesure, d'un jeu de carte. Il distingue le concept d'ordre, référant à un « état idéal » où il y a un accord complet à une règle, et le concept d'ordonnement (« *orderliness* ») comme mesure

de l'écart entre une « entité réel » et cet état idéal. Cette distinction peut être utile en autant que la règle définissant cet état idéal soit justifiée et que soit aussi justifié cet état idéal, s'il est entendu comme une norme ou un objectif. Mais ce n'est clairement pas toujours le cas. Cette règle définissant l'ordre est généralement implicite et l'ordre est associé à la singularité d'un état parmi une gamme de possibles. Par exemple, un état serait ordonné *parce qu'il* ne pourrait être réalisé que d'une seule façon (Müller & Weiss 2008). Toutefois, il s'agit d'une condition sans doute nécessaire à la notion d'ordre, mais elle demeure insuffisante : plusieurs états désordonnés sont uniques et ils peuvent correspondre à une entropie élevée ou non.

Ce caractère vague du concept de désordre favorise les associations plus ou moins rigoureuses avec d'autres concepts, comme celui, déjà mentionné, de « désorganisation ». Le concept de désorganisation, et son opposé, celui d'organisation, demeurent souvent indéfinis dans la plupart des discussions abordant des sujets comme les systèmes complexes, l'émergence de la vie et les processus loin de l'équilibre thermodynamique ; l'article de Wimsatt (1972) et celui de Bickhard (2011) en sont des exemples³¹. Denbigh (1989a), reconnaissant la difficulté de définir ce concept, argue que les concepts d'ordre et d'organisation ne sont pas synonymes en donnant l'exemple d'un papier peint à motifs et celui d'un tableau de Cézanne, le premier étant plus « ordonnancé » (« *orderly* ») mais moins « organisé » que le second. On conçoit donc que le concept de désorganisation est largement entendu ici comme un concept primitif appréhendé de manière intuitive. Il est vrai cependant que l'interprétation d'un concept par un autre ne peut procéder par régression à l'infini mais doit plutôt s'appuyer sur des concepts bien définis, nominalement ou empiriquement. L'exemple précédent, quoiqu'illustratif davantage que démonstratif, semble s'appuyer sur une définition de l'organisation faisant intervenir des concepts avec lesquels on l'associe d'ordinaire (dans le dictionnaire par exemple), comme celui de « structure » et de « mode de fonctionnement ». La définition qu'offre Bunge (2003 :19), comme « collection de relations parmi les composants d'un système », semble adéquate ici puisqu'elle a le mérite d'être claire et de s'appuyer sur des concepts génériques et primitifs, comme celui de « collection », de « relation » et de « système ». On comprend ainsi mieux l'exemple où le papier peint présente un petit nombre de relations parmi ses composants, donc une petite collection de relations, tandis que le tableau présente un plus grand nombre de relations et donc une plus grande collection de relations. Mais défini de la sorte le concept d'organisation semble se placer sur un

³¹ Bickhard défend une métaphysique des processus où les stabilités des « processus d'organisation » sont le point focal de ce qui doit être expliqué dans les phénomènes naturels, sans jamais toutefois définir ce qu'il entend par « organisation » ou « processus d'organisation » ; il affirme : « *Everything is stable (or not so stable) organizations of processes.* » (Bickhard 2011 : 97) Wimsatt (1972 : 76) affirme : « *one would expect that at least the vast majority of complex evolved systems would be hierarchially organized* » ; mais il ne définit pas l'organisation, quoiqu'il laisse entendre qu'elle covarie avec la complexité : « *Unless causal relations or state variables were organized in a rather specific way relative to the different theoretical perspectives, one would expect systems to be interactionally complex.* » (75)

continuum avec celui d'ordre, où l'un exprimerait la même idée que le second mais avec plus d'intensité. Ainsi, un système « organisé » serait « très ordonné » et tout système « organisé » serait « ordonné », mais l'inverse ne serait pas vrai.

Outre l'objection que l'on peut adresser à la notion même de désordre, l'interprétation de l'entropie comme « mesure du désordre » est l'objet de critiques visant surtout l'inadéquation au concept d'entropie, comme quoi la classe des systèmes d'entropie élevée et celle des systèmes désordonnés et/ou désorganisés ne seraient pas coextensives, et de même entre les systèmes de faible entropie et les systèmes ordonnés et/ou organisés. Les arguments visent donc à montrer que l'association entre deux concepts, où l'un permet d'interpréter l'autre, n'est pas toujours valide car s'y présentent des exceptions. Ce type d'arguments visant à contrer un énoncé universel s'appuie sur des contre-exemples, et Klein (1953) en présente quelques-uns. L'un d'eux consiste en deux moles de chlorure d'hydrogène, tandis que l'autre est une mixture d'une mole d'hydrogène et une mole de chlorure, le premier système étant, du moins intuitivement, plus « organisé » que le second, bien qu'un calcul simple montre que l'entropie du premier est plus élevée que celle du second système. Un autre exemple est celui de deux systèmes dont la température est au zéro absolu et dont les entropies respectives, en vertu du troisième principe de la thermodynamique, sont toutes deux nulles ; or il se trouve, bien entendu, que plusieurs systèmes au zéro absolu peuvent présenter des degrés d'organisation fort différents. Puisque, selon lui, le « degré d'ordre » serait mesuré par l'entropie, il en conclut que le « degré d'ordre » n'est pas la même chose que le « degré d'organisation ». Mais ces discussions n'apportent rien si la validité de la métaphore est postulée et si on ne parvient pas à clarifier « ordre » et « organisation ». Il semble ici que ce soit le concept d'entropie qui vienne éclairer celui d'ordre plutôt que l'inverse.

Un autre type d'objection consiste à établir l'échec du concept de désordre, ou d'autres concepts jugés équivalents, à satisfaire les propriétés attribuées au concept d'entropie. Il s'agit de ces contraintes logiques dont parle Russell (1948), discutées précédemment, à condition que le concept de désordre utilisé soit bien défini. L'un de ces arguments identifie la propriété de l'extensivité de l'entropie (Landsberg 1985 ; Leff 2007 ; Ben-Naim 2008 ; Michaelides 2008). Comme l'entropie est une propriété extensive (Chapitre 4), l'entropie d'un système double si, toute chose étant égale par ailleurs, son volume double. Mais il est loin d'être évident que le désordre double aussi si toutes les variables du système demeurent les mêmes. Par exemple, est-ce que le désordre d'une salle de classe est deux fois plus élevé que, toute chose étant égale par ailleurs, le désordre d'une salle de classe à moitié moins grande ? Ce serait ainsi le désordre qui est une propriété intensive et ce serait plutôt l'entropie spécifique, l'entropie par unité de masse ou de volume, qui est extensive, qui peut être légitimement associée au désordre.

Mais si l'objectif initial est d'établir un équivalent mécanique du concept d'entropie thermodynamique, alors il semble que ce soit bien au sein de cette théorie qu'il faille chercher l'interprétation. Or aucune définition du désordre n'est fournie en thermodynamique. La question de la légitimité de l'interprétation du désordre touche donc aussi bien la définition et la signification de ce concept que la théorie ou la discipline employée pour établir cette définition. Puisque la théorie de la thermodynamique est muette à ce propos, elle ne privilégie pas, *a priori*, une interprétation du concept d'entropie statistique en termes de « désordre ». Outre l'objection à l'effet que la notion de désordre soit vague et subjective puisqu'elle semble n'être appréhendée que par des impressions, cette métaphore à la base de l'interprétation n'embrasse pas toutes les situations où s'applique le concept d'entropie et échoue dans certains cas à représenter certaines de ses propriétés, comme l'extensivité.

5.3.4.3 IGNORANCE ET INFORMATION

L'interprétation de l'ignorance et celle de l'information (distinguées plus loin) découle de l'expérience de pensée de Maxwell, une « *gedankenexperiment* » (Chapitre 3), qui a introduit un « démon », un être avec des capacités sensorielles extraordinaires pouvant « voir » la vitesse des particules dans une enceinte partagée par une paroi munie d'une trappe qu'il contrôle sans frottement, en sorte qu'il est en mesure de séparer les particules selon leur vitesse, et donc de produire une différence de température entre les deux sections de l'enceinte sans dépense de travail, violant ainsi le second principe de la thermodynamique et (donc ?) le principe entropique. Autrement dit, cette différence de température pourrait être utilisée pour produire un mouvement perpétuel de second type. La conclusion qui semble s'imposer est que ce principe tient à notre incapacité à déterminer l'état microscopique mécanique d'un grand nombre de particules, à notre *ignorance* de cet état, à notre *manque d'information* (discuté plus loin). Car, semble-t-il (Albert 2000 : 39), ce n'est que parce que ce démon est en mesure d'établir *plus* que ce qui est exprimé par la description thermodynamique du gaz, parce qu'il est en mesure d'établir son état microscopique, ou microcondition, qu'il peut provoquer une violation du second principe. Un tel principe relatif à nos capacités cognitives ne serait pas un « vrai » principe portant sur la « nature » des choses.

La littérature sur le sujet, tant en physique qu'en philosophie, est impressionnante³² (voir Earman & Norton 1998 et 1999 ; Bub 2002 ; Leff & Rex 2003 ; Maroney 2009), au point où l'on semble avoir perdu de vue le besoin, ou non, d'un « exorcisme du démon » (Earman & Norton 1998 et 1999). La réelle portée de cette expérience de pensée pour le concept d'entropie n'est pas non plus évidente

³² Dans la seconde édition de leur ouvrage compilant les diverses discussions sur le démon de Maxwell, Leff & Rex (2003 : xii) avancent qu'il y a eu plus de références générées en 13 ans depuis la première édition que depuis la première apparition de cette expérience de pensée dans la littérature dans les années 1870.

puisque les discussions s'y référant adoptent différents points de vue quant au statut du second principe de la thermodynamique et du principe entropique, à savoir si leur validité respective doit être tenue pour acquise ou non. Si, par exemple, cette validité est postulée, alors une éventuelle violation, par expérience de pensée toujours, ne peut mener qu'à la conclusion voulant que ce genre de démon est physiquement impossible. En revanche, certains auteurs, comme Popper (1957), ont proposés des versions modifiées du principe entropique. Si bien que l'on ne semble plus savoir qui ou quoi « exorciser ».

Sans doute, le principal problème avec ce genre d'expérience de pensée concerne le bon usage des idéalizations et des fictions : lesquelles sont valables et dans quelles limites ? Dans l'exemple du démon de Maxwell, l'une de ces idéalizations consiste en une trappe se déplaçant sans frottement. On peut justifier son emploi en soutenant que l'effet du frottement est négligeable, que l'énergie impliquée dans le frottement, par exemple, est de beaucoup moindre à l'énergie impliquée dans le processus total qui, en définitive, se conclut par une augmentation de la différence de température sans dépense de travail. Or, une telle justification semble difficile à défendre compte tenu du fait que l'énergie cinétique d'une particule microscopique ne semble pas, *prima facie*, largement inférieure à l'énergie nécessaire au déplacement, et au frottement qu'il implique, d'une trappe composée de plusieurs particules. Une autre idéalisation concerne l'absence de « coût » associé à la mesure, si tant est qu'on le considère comme un processus physique, permettant au démon de recueillir de l'information sur le micro-état du système. L'énergie impliquée dans le processus visant à voir les particules, à déterminer leur état dynamique, n'est généralement pas prise en compte. Il semble plausible que ce coût énergétique soit suffisant pour troubler, empêcher la construction d'un mouvement perpétuel, bref à exorciser le démon. On pourrait arguer en outre que la mesure implique des photons qui, à cette échelle, font intervenir le principe d'indétermination d'Heisenberg, de sorte que le démon pourrait déterminer si une particule est rapide ou non, mais qu'il ne pourrait pas savoir où elle se trouve, et vice-versa³³.

Outre la question du traitement adéquat des idéalizations, il y a celle, non moins importante, de la définition du système. Car, comme on l'a vu (Chapitre 4), puisque l'entropie peut diminuer localement, si le système n'est pas défini, circonscrit correctement, les inférences sont biaisées. Ainsi, du moment où le démon est aussi considéré comme un système thermodynamique obéissant au second principe, c'est-à-dire que son état doit être considéré afin déterminer si oui ou non il y a violation du second principe, il importe d'en déterminer les conséquences pour le processus total et donc pour la conclusion quant à

³³ Cependant, ce genre d'argument nécessite sans doute un *autre* concept d'entropie, quantique cette fois. Quoiqu'il en soit, la définition de la version quantique du concept d'entropie pose des problèmes qui remettent en doute sa compatibilité avec l'impossibilité du mouvement perpétuel (voir section 5.4.2).

la soi-disant violation. Ce qui veut dire, plus généralement, que si le processus de mesure modifie par une quelconque interaction l'état du démon, alors il faut en tenir compte dans le bilan des transformations énergétiques. Par exemple, on ne peut conclure à la violation du principe entropique sur la simple base qu'il y a diminution d'entropie dans ladite enceinte lorsqu'il y a interaction avec le démon, car celui-ci pourrait bien surcompenser cette diminution d'entropie et en conséquence « sauver » ce principe. Si tel est le cas, il faut établir les propriétés du démon, comme système physique, susceptible d'influencer la valeur de l'entropie totale du système composé de l'enceinte et du démon. En général, les discussions à ce sujet se concentrent sur le processus de « mesure de l'information », le coût entropique de cette mesure, déterminant l'état microscopique du gaz, et si ce processus nécessite ou non un être intelligent comme est censé l'être le démon maxwellien. Plus particulièrement, elles s'articulent autour du « coût de l'information », à savoir si un gain ou une perte d'information implique une diminution ou une augmentation d'entropie.

Il y a deux problèmes majeurs, à mon sens, qui limitent sérieusement un quelconque progrès en la matière. D'une part, et cela est en lien avec le problème exposé plus tôt de la définition du système, les diverses descriptions des transformations énergétiques impliquées dans ces expériences de pensée, comme le travail fourni au gaz par une compression ou encore la libération de chaleur dans un réservoir thermique entourant l'enceinte, sont souvent mal engagées. C'est-à-dire que les bilans énergétiques sur lesquels s'appuient les conclusions quant à la violation ou non du principe entropique ne sont pas toujours effectués avec rigueur et cohérence tout au long du raisonnement. Par exemple, en lien avec ce qui précède, si les idéalizations sont inadéquates, plus spécifiquement si la friction est négligée alors qu'elle représente, en principe, un phénomène dont l'influence est suffisamment importante, alors les conclusions peuvent être significativement biaisées³⁴. Ou encore, si l'intervention du démon visant à mesurer, à déterminer l'état microscopique du gaz est bien un processus physique dont il faut tenir compte, susceptible de modifier justement cet état de manière significative, alors son abstraction du processus serait injustifiée. D'autre part, et sans doute plus grave encore, les argumentations employées font souvent appel à trois versions différentes du concept d'entropie : l'entropie thermodynamique afin de traiter des opérations empruntées aux manipulations macroscopiques, comme une compression isothermique, l'entropie statistique sans distinction claire entre ses différentes versions, l'entropie informationnelle afin de tenir compte des gains et pertes

³⁴ Voilà un problème d'importance. Et on notera son caractère normatif puisqu'il nécessite une détermination de ce qu'il faut faire, de ce qui est *suffisamment important*, etc. Ce problème est celui, justement, du bon usage des idéalizations. Ashby & Miller (1970 : 31) donne l'exemple d'un disque tournant sur un axe : la densité de probabilité qu'il s'arrête à un certain angle est indépendante de la variable continue qu'est l'angle si la friction peut être négligée, plus précisément s'il n'y a pas plus de friction à un endroit plutôt qu'à un autre autour de cet axe. Alors, dans ce cas, le caractère normatif est déterminé par une indépendance statistique. Mais cette indépendance va être justifiée si la fréquence relative des angles obtenus est uniforme sur plusieurs essais. (Voir aussi Batterman 2005 : 235.)

d'information, et parfois aussi l'entropie quantique. Il n'est donc pas clair quel concept d'entropie doit être utilisé, ni si l'emploi de plusieurs d'entre eux est vraiment cohérent, et par conséquent quelle portée peuvent avoir les conclusions de ces expériences de pensée.

À supposer que ces ambiguïtés puissent être surmontées, il faudrait tout de même, pour que ces discussions soient épistémologiquement pertinentes et intéressantes, que l'on précise ce que l'on entend par une « violation » du principe entropique ou du second principe. Car si le démon de Maxwell est bien un démon au sens d'une « créature surnaturelle », si on lui attribue des pouvoirs transcendant les lois de la physique, alors il y a lieu d'être dubitatif quant aux conclusions physiques que l'on puisse en tirer. Certains ont même avancé (Smoluchowski, in Earman & Norton 1998) que l'intelligence, en tant que phénomène, se situait au-delà des lois de la physique. Cela laisse entendre deux choses. D'abord que le phénomène de l'intelligence est bien à l'œuvre dans le processus décrit dans cette expérience de pensée, mais il n'est pas évident que la détection de la vitesse ou de la position d'une particule constitue vraiment un cas d'intelligence comme on l'entend habituellement, impliquant des capacités d'apprentissage, de jugement et de compréhension. Ensuite, quoique de manière plutôt ambiguë, on laisse entendre que soit une expérience de pensée avec de tels démons est intrinsèquement incohérente ou simplement mal conçue en négligeant d'autres phénomènes qui exigeraient un traitement « extra-physique », qui seraient biologiquement irréductibles par exemple, soit le processus physique total doit prendre en compte ce phénomène de l'intelligence impliqué dans l'acquisition d'information par ledit démon, c'est-à-dire en le « naturalisant ». Comme le démon représente le système chargé de l'acquisition d'information sur l'état dynamique du gaz, cette naturalisation revient à considérer cette acquisition comme un processus physique susceptible d'être soumis aux lois de la thermodynamique (discuté plus loin). Il faut donc voir le choix qui s'impose ici : soit il s'agit d'une créature transcendant les lois de la physique, auquel cas celles-ci sont violées *ex hypothesi* par son intervention, soit il s'agit d'un système physique et donc, *ex hypothesi* aussi, il devrait être soumis aux lois de la thermodynamique, ne leur apportant ni preuve ni réfutation.

Mais l'intervention d'une telle créature n'est pas nécessaire, et cela a déjà été noté. Ainsi, il est possible, quoique très peu probable, comme l'a déjà souligné Boltzmann et selon les termes déjà discutés, qu'un gaz se retrouve de lui-même, sans intervention extérieure, dans un volume plus petit que son volume initial, ou encore, dit autrement, qu'un piston soit comprimé par la seule action du gaz qui l'entoure. Il s'agit de deux cas de phénomènes en contradiction avec le principe entropique, exprimé dans sa forme absolue, thermodynamique, et non probabiliste. Pour reprendre les mots de Earman & Norton (1998a), il s'agit d'une « violation directe » (« *straight violation* »). Mais ils mentionnent aussi une « violation enjolivée » (« *embellished violation* ») qui implique la possibilité d'exploiter cette diminution d'entropie pour produire un travail. Deux points sont à préciser : d'abord, il faut dire si

cette production de travail est suffisamment grande, ou effective « en moyenne », de manière « continue », sur une longue période, etc., ou si cette conversion de la chaleur en travail n'est que temporaire, sans véritable moyen de l'exploiter de manière utile ; ensuite, une condition souventes fois mentionnée pour une telle production concerne certaines capacités cognitives, comme chez le démon, ou un dispositif capable mesurer et d'acquérir de l'information sur un système. Ces capacités cognitives ou informationnelles sont nécessaires, semble-t-il, afin d'éviter les conséquences probabilistes de la version statistique du principe entropique. Car, si le principe entropique est vraiment de nature probabiliste, comme quoi il est *très* probable qu'un système isolé soit dans un état évoluant vers un état d'entropie supérieure, alors il n'est évidemment pas *impossible* que l'entropie de ce système diminue. Seulement, compte tenu du calcul de ces probabilités et toute chose étant égale par ailleurs, le temps pour que cela se produise risque d'être de l'ordre de l'âge de l'univers, un « peu de patience » selon l'euphémisme de Poincaré. Ainsi, l'information serait nécessaire pour aller « un peu plus vite ». C'est ainsi que le concept d'entropie s'offre un lien, dans l'une de ses formes, avec celui d'information, un lien révélant la « nature statistique » (Maxwell 1878, in 1890b : 670) du principe entropique, mais aussi, on le verra, ce qu'on pourrait appeler l'aspect physique du processus d'acquisition d'information.

5.3.4.3.1 Machine de Szilard et principe de Landauer

La *machine de Szilard* est justement une expérience pensée de Leo Szilard (1929) visant à déterminer les conditions permettant la construction d'une machine à mouvement perpétuel de second type opérée par un « être intelligent », comme dans l'expérience de pensée de Maxwell. Cet appareil, un genre de petite machine à vapeur, consiste en un cylindre séparé par une partition jouant le rôle d'un piston, à l'intérieur duquel se trouve une seule particule et qui reçoit de la chaleur d'un réservoir thermique. Si la localisation de la particule est connue, grâce à une mesure par exemple, alors il est possible d'harnacher le piston-partition, qui se déplace en raison de la pression exercée par la particule, à un poids et ainsi produire, *en moyenne*, un travail. Il faut alors préciser une chose ici, car il peut sembler superflu d'exiger une étape de mesure puisqu'il est possible de placer le poids, « par chance », du « bon » côté et ainsi produire un travail. C'est précisément ici que la discussion se complique et que s'exposent les difficultés associées aux dialogues entre les descriptions du « niveau microscopique » et du « niveau macroscopique ». En effet, il n'est pas suffisant, au niveau macroscopique, de produire un travail « par chance », au gré des fluctuations moléculaires aléatoires. Il y a une notion de contrôle liée à la celle de travail qui fait en sorte que celui-ci doit être « reproductible », de manière « continue », etc., ce qu'est censé garantir la mesure³⁵. Pour que les conséquences thermodynamiques soient pertinentes, il faut

³⁵ Je ne crois pas avoir besoin de plus de mots pour exposer l'ampleur des questions épistémologiques que tout cela soulève.

donc une information fiable et détaillée sur l'état du système et une intervention en fonction de cette information. Il y a évidemment plusieurs idéalizations qui entrent en ligne de compte. Mise à part celle négligeant certains phénomènes plutôt exotiques qui pourraient intervenir à l'échelle d'une seule particule, comme l'effet Casimir³⁶, l'on suppose un processus quasi-statique, une expansion isothermique réversible, le maintien de l'énergie cinétique de la particule, et que le travail produit provient du transfert de chaleur du réservoir thermique.

En fait, comme il n'y a pas de distinction claire, au niveau microscopique, entre la chaleur et le travail, on peut être dubitatif quant au recours aux intuitions dérivées de la thermodynamique. Ainsi, le réservoir transfère de l'énergie prétendument sous forme de chaleur à la particule, qui la transfère, cette fois sous forme de travail, du moins par hypothèse, au piston-partition. Cette conversion complète de la chaleur (Q) en travail (W) contredit bien sûr le second principe de la thermodynamique. Car il a été vu (Chapitre 4) qu'une production de travail à partir d'une quantité de chaleur exige un transfert d'entropie à l'extérieur du système, dans son environnement, d'une quantité supérieure à la diminution d'entropie à l'intérieur du système ou dans le réservoir thermique en raison du transfert de chaleur. Par exemple, dans leur présentation de la machine de Szilard, Leff & Rex (2003 : 14) avancent que la particule reçoit une « énergie $Q = W$ fournie par la chaleur », ce qui signifie une conversion complète. Si les prémisses contredisent ce principe, il ne faut pas s'étonner que la conclusion statue en faveur d'une violation dudit principe. Je poursuis maintenant la discussion avec la suite du raisonnement de Szilard.

Selon ce dernier, le second principe et le principe entropique (il ne les distingue pas) peuvent être sauvés, l'« exorcisme » est possible, si, comme je l'ai mentionné, on inclut ce démon dans le processus physique d'acquisition d'information, bref en le naturalisant. Cependant, l'exorcisme du démon maxwellien est mitigé, ou plutôt éludé, car il procède du « postulat qu'une compensation complète [d'entropie] est faite selon le second principe [principe entropique] » (Szilard 1929 : 124). En somme, cette compensation est déterminée à partir de la description du processus physique de la machine décrite précédemment et du postulat que le principe entropique est valide et s'applique à ce processus. Ce n'est donc pas une preuve de ce principe. Le raisonnement est plutôt le suivant : un système physique, la machine, est en interaction avec un autre système physique, l'être intelligent ; puisqu'il manque, dans la description du fonctionnement de la machine, une certaine quantité physique pour satisfaire à un principe dont la validité est postulée, il faut que cette quantité se trouve dans le système

³⁶ Les fluctuations quantiques créent des forces impactant les corps macroscopiques. Ainsi, ces fluctuations exercent sur des plaques parallèles une pression, une force attractive entre ces plaques, non négligeable à l'échelle nanométrique. Mais des travaux récents montrent que cette force peut être répulsive lorsque les matériaux sont adéquatement choisis (Munday, Capasso & Parsegian 2009).

avec lequel il y a interaction ; donc, il y a compensation dans l'être intelligent qui procède aux mesures, à l'acquisition d'information. En se basant sur la loi des gaz parfaits, $PV = \mathcal{N}kT$, dans le cas où $\mathcal{N} = 1$, on peut déterminer le travail maximum extrait de cet arrangement :

$$W = \int PdV = \int (1/kT)V^{-1} dV = kT \ln(V)_{[V, V/2]} = kT \ln(V/2) - \ln(V) = kT \ln(V/2V) = -kT \ln 2.$$

Cette quantité est donc interprétée comme la quantité de travail nécessaire (W) à l'acquisition d'information lors du processus de mesure par le démon.

Ce résultat a été très tôt interprété comme déterminant le « coût de l'information », comme quoi « $kT \ln 2$ » représenterait la « valeur thermodynamique de la connaissance » (von Neuman 1955 : 400). Et ce, avant même que les travaux de Shannon établissent la version informationnelle du concept d'entropie. Ainsi, plusieurs travaux, dont ceux de von Neumann, de Brillouin, de Rothstein, et de Gabor, ont développé l'idée que l'acquisition d'information par une mesure requiert la dissipation d'une quantité d'énergie équivalente à $kT \ln 2$ pour chaque bit d'information acquis. À partir de ce moment les considérations thermodynamiques ont plus souvent qu'autrement été éloignées, insistant plutôt sur une « mesure statistique de notre ignorance » ou sur un « manque d'information », voire même, comme le déplore Gyftopoulos (2002 : 411), « comme un concept subjectif plutôt que physique ». Avant de discuter cette autre association avec la notion d'information, je me tourne vers la discussion à propos du processus physique d'acquisition d'information, que l'on retrouve typiquement en informatique.

Associant l'« irréversibilité logique » à l'« irréversibilité physique », Landauer (1961) tente d'établir quelles sont les limitations physiques de la vitesse et de la compacité d'un appareil informatique. Il s'appuie d'abord sur le raisonnement de Brillouin à l'effet qu'un processus de mesure demande une dissipation d'énergie de l'ordre de kT . Il soutient ensuite qu'un processus de mesure est analogue à un processus computationnel dont les états logiques doivent correspondre à des états physiques. Il distingue alors les opérations logiquement réversibles des opérations logiquement irréversibles : une opération est logiquement réversible si l'état intrant (*input*) peut être correctement distingué de l'état extrant (*output*). Ainsi, puisqu'une opération logiquement réversible consiste en une application bijective ($1 : 1$), il s'ensuit qu'une opération logiquement irréversible réduit l'espace des phases logique et donc, *ex hypothesi*, comprime l'espace des phases physique. Landauer soutient que cette compression doit être accompagnée d'un accroissement d'entropie dans l'environnement sous la forme d'une dissipation de chaleur. Pour quantifier la dissipation de chaleur, il utilise l'opération logique irréversible la plus simple qu'il considère être la réinitialisation d'un bit. Cette opération implique deux états logiques comme intrants (conventionnellement ZERO et UN) et rend toujours ZERO comme extrant pour la réinitialisation l'état logique (on peut aussi prendre UN par convention). Il est commun de réaliser cette

opération avec un dispositif similaire à la machine Szilard, où la partie gauche de la boîte correspond à l'état ZERO. La réinitialisation du bit exige que la particule soit située dans la partie gauche, ce qui demande un travail de $kT \ln 2$. Une version du *principe de Landauer* stipule qu'il n'y a pas de réalisation physiquement possible de l'opération de réinitialisation pouvant faire mieux qu'une réinitialisation convertissant moins qu'une quantité $kT \ln 2$ de travail en chaleur. Bien que Landauer réfère à cette opération comme une *réinitialisation*, la littérature subséquente a plutôt référé à un *effacement*. Mais il peut facilement y avoir confusion puisque le terme « effacement » semble signifier une opération où l'information est simplement détruite sans que le système retourne à un état déterminé à ZERO.

Les débats quant à la validité de ce principe sont tout à fait d'actualité : Earman & Norton (1999) soutiennent que les hypothèses de Landauer sont controversées ; Maroney (2005) que la réversibilité logique n'implique pas et n'est pas impliquée par la réversibilité thermodynamique ; Norton (2005) soutient que la thèse à l'effet que l'effacement comprime l'espace des phases est erronée et qu'en conséquence le principe n'est pas encore démontré ; Ladyman *et al.* (2007), tenant pour valide le second principe de la thermodynamique, propose une « nouvelle preuve » de ce principe de Landauer ; Norton (2011) argue que cette preuve est intenable ; Ladyman & Robertson (2013) la défendent ; et cela se poursuit encore (Norton 2013 ; Ladyman & Robertson 2014). De plus, deux articles récents parus dans *Nature* semblent corroborer le principe de Landauer (Bérut *et al.* 2012 ; Roldàn *et al.* 2014). Ces discussions se penchent principalement sur les notions d'entropie et d'information comme processus physiques, impliquées dans des descriptions causales d'interventions sur un système physique. Malgré les problèmes auxquelles elles font face, avec plus ou moins de succès, elles s'appuient sur une approche, sur des intuitions thermodynamiques. C'est-à-dire qu'une mesure empirique est considérée comme impliquant une augmentation ou une diminution d'entropie thermodynamique, aussi corrélée avec un gain ou une perte d'information. En effet, certains établissent cette corrélation avec l'acquisition d'information, d'autres avec son effacement. C'est ce que j'appellerais l'approche *substantielle* de la connexion entre l'entropie et l'information. Il s'agit d'une interprétation de l'information en termes d'entropie plutôt que l'inverse, où l'échange d'information est considéré du point de vue physique.

Mais il y a aussi une approche *sémantique* de cette connexion, très influente, établissant, pour certains, une « identité » (Ben-Naim 2008 : 22) entre le concept d'entropie thermodynamique et celui d'information au sens qu'en donne Shannon. Selon Brillouin (1956), l'entropie et la quantité d'information pourraient être considérées comme deux faces complémentaires d'une même grandeur. Ainsi, le « manque d'information » (ou l'ignorance) correspondrait à l'entropie et l'« information » à la

« néguentropie » (entropie négative). Cette approche définit une quantité qui semble être « abstraite du monde dans lequel elle est intégrée »³⁷ (Keller 2011 : 175). Avant d'en discuter, on se rappellera que Shannon a procédé à une analogie formelle entre sa définition de l'entropie informationnelle et l'entropie statistique. Il faut rappeler également que cette approche oblitère parfois l'autre approche en faisant abstraction du processus physique nécessaire pour obtenir de l'information sur un système réel. Cela deviendra encore plus évident, à tout le moins son importance sera mise en exergue, dans notre discussion sur l'interprétation en termes de travail (section 5.3.4.4) et sur l'entropie quantique (section 5.4.2).

5.3.4.3.2 Théorie de l'information

En *théorie de l'information*, trois concepts clés se partagent l'essentiel de son contenu, soit les concepts d'émetteur, de récepteur et de message. L'accent est alors placé sur le degré d'incertitude que le receveur peut avoir sur le message qui arrive. C'est pourquoi les probabilités p_i sont généralement interprétées comme des probabilités épistémiques, des degrés de croyances. Cette interprétation est néanmoins tributaire de la source de l'attribution des probabilités aux différents messages. Si l'attribution découle de et exprime les croyances des receveurs, alors, compte tenu de la diversité de ces croyances, les probabilités seront aussi très différentes d'un receveur à l'autre. Les probabilités sont ainsi subjectives au sens où elles ne sont pas indépendantes des croyances des sujets, des agents que sont les receveurs. En revanche, si l'attribution des probabilités est indépendante des receveurs, alors celles-ci peuvent être considérées comme objectives. C'est le cas, par exemple, de l'attribution se faisant sur la base d'un modèle, d'une idéalisation comme celle d'un « agent rationnel » où l'attribution des probabilités est basée sur une règle voulant que l'entropie de Shannon $S_{s,d}(P)$ doit être maximisée, ce qui sélectionne une distribution unique. Il s'agit du *principe de maximisation de l'entropie* de Jaynes. Par ailleurs, l'attribution des probabilités peut aussi être indépendante, dans une certaine mesure, des agents, si l'exigence est placée sur l'adéquation de ces probabilités avec les probabilités d'occurrence de l'événement auquel elle réfère, donc avec des probabilités ontiques.

L'objectif de Jaynes était de montrer que le lien entre entropie et information était fondamental et que les principales lois de la mécanique statistique se présentaient naturellement, ou « émergeaient » (Duncan & Semura 2007 : 1769), si l'on traitait notre ignorance des états microscopiques de manière non-biaisée. Ainsi, son principe de maximisation de l'entropie serait un outil pour faire les meilleures inférences possibles compte tenu d'une information incomplète. La mécanique statistique ne serait qu'un cas particulier de cette méthode. Cette manière non-biaisée d'attribution de valeur de

³⁷ « [...] *abstracted away from the material world in which it is embedded* ».

probabilités consiste en ce critère formel du principe de maximisation de l'entropie, une sorte de « généralisation du fameux principe d'indifférence » (Sklar 1993 : 190). En ce sens, cette méthode d'inférence est à la fois subjective et objective en ceci qu'elle réfère à des degrés de croyances, et c'est pourquoi on la qualifie de « subjectiviste » (Bunge 2010), mais elle propose aussi un critère claire encadrant les inférences dans des cas de sous-détermination de systèmes physiques, et c'est pourquoi on la classe sous le « bayesianisme objectif » (Sklar 1993 : 190). Encore une fois (voir section 5.1.3), cette dualité est irréductible puisque, bien que les probabilités ne soient pas observables à proprement dit, elles sont ici paramétrées empiriquement ; mais aussi parce qu'elles s'appliquent à un système nomologiquement déterministe mais néanmoins sous-déterminé, c'est-à-dire dont la connaissance est par hypothèse incomplète.

La méthode de Jaynes apparaît toutefois comme une méthode formelle d'inférence infalsifiable. Cela n'en fait pas une méthode invalide ou inutile, même que le principe de maximisation de l'entropie est employé avec succès dans la détermination des densités spectrales, mais cela ne définit pas une quantité, susceptible d'être mesurée empiriquement, qui augmente dans le temps, de manière irréversible. La question est la suivante : comment un concept comme l'entropie thermodynamique, défini par celui de la chaleur et celui de la température et présentant donc les unités de l'énergie divisée par la température absolue, peut-il être « réduit » à celui d'information sans unité physique ? L'aspect particulièrement épistémique convoyé par des notions comme celles d'information mais aussi de probabilité s'introduit presque naturellement, intuitivement dans la caractérisation *partielle* d'un système physique, en ce sens où une telle caractérisation ne peut que renvoyer à nos connaissances, toutes aussi partielles, du système.

Si cette caractérisation était complète et exhaustive, alors il y aurait en quelque sorte superposition, coextensivité entre celle du référent et celle de nos connaissances du référent. Pour l'exprimer autrement, l'énoncé « la particule est à x » et l'énoncé « la particule est à x_1 ou x_2 ou x_3 » n'ont pas la même sémantique puisque l'énoncé disjonctif exprime des possibilités et le sujet semble référer à nos connaissances, comme quoi le sujet du premier énoncé peut recevoir une définition ostensible mais pas le second. Je ne souhaite pas ici m'avancer plus loin dans ces considérations linguistiques, mais simplement insister sur le fait qu'il est naturel d'associer une sous-détermination empirique avec des connaissances incomplètes et ces dernières à quelque chose de « subjectif ». Toutefois, je puis savoir, objectivement, qu'une particule se trouve dans une boîte sans savoir où elle se trouve : cette incertitude n'a rien de subjectif, mais la façon de la traiter peut être arbitraire.

En conclusion, le concept d'entropie entretient des liens avec celui d'information mais ne peut s'y réduire ni l'englober. Des résultats empiriques récents (Bérut *et al.* 2012 ; Roldàn *et al.* 2014) rappellent

ce lien. L'idée est féconde mais elle peut « donner lieu à des généralisations ou assimilations imprudentes » (Monod 1970 : 212). De fait, comme les processus d'acquisition et d'effacement de l'information sont des processus physiques, ils doivent être soumis aux lois de la physique comme le principe entropique. Il s'agit de l'approche substantielle exemplifiant le lien entre entropie et information³⁸. Elle implique une intervention de l'observateur, soit un processus physique, et c'est pourquoi l'entropie ne saurait se dissoudre dans les capacités cognitives de cet observateur. Dans l'approche sémantique, il est plutôt question du contenu en information d'une description d'un système physique, d'un message. À défaut d'un principe clair circonscrivant le comportement d'une éventuelle quantité bien définie que serait l'information, comme un principe stipulant sa conservation, on est en droit de s'attendre à ce que son traitement dans cette approche demeure métaphorique, avec les dangers épistémiques que cela comporte.

5.3.4.4 TRAVAIL

L'interprétation du concept d'entropie comme celle de la capacité à effectuer un travail provient du second principe de la thermodynamique stipulant l'impossibilité pour un système donné d'une conversion complète de la chaleur en travail. Elle s'appuie donc sur les bases solides du principe de l'impossibilité du mouvement perpétuel de second type et elle a une histoire aussi longue que la thermodynamique. L'idée générale est bien résumée par Brunhes (1909 : 400, italiques originaux) :

Quand on a laissé l'équilibre thermique s'établir dans un système matériel, il a [...] perdu quelque chose, mais il est une chose qu'il a conservée. Il n'a plus la même *capacité à produire un travail*. Mais il a gardé la même *énergie*. Si elle n'est plus à notre portée, elle existe toujours. Seulement il n'est plus possible de la récupérer sans dépense.

Ce que le système a perdu serait non pas une *quantité* d'énergie mais une *qualité* d'énergie. En ce sens l'augmentation d'entropie correspondrait à une « dégradation d'énergie ». Avec un vocabulaire plus archaïque quoique philosophiquement connoté, on peut citer Ronsard selon lequel « la matière demeure et la forme se perd ». (Ce vers permet peut-être d'appréhender la précédente distinction, mais en prendre la pleine mesure nécessiterait un détour significatif vers les subtilités introduites par la physique contemporaine, notamment avec la théorie de la relativité, où la notion d'énergie reçoit des caractérisations exigeant un examen approfondi de notions fondamentales comme celles de « propriété », d'« état » et de « processus ».)

³⁸ Je crois qu'il s'agit d'un terrain d'analyse important pour le problème de la mesure en général et surtout en mécanique quantique (section 5.4.2).

En caractérisant l'entropie par une fonction d'état, cette interprétation était vouée à des difficultés sans doute insurmontables puisque le travail n'est *pas* une fonction d'état. Ainsi, l'entropie, dans sa version thermodynamique mais aussi statistique, est une fonction d'état définie par certains paramètres. Si, par exemple, on double le nombre de molécules d'un système, doublant sa masse, alors son entropie statistique double, mais il est loin d'être évident que sa capacité à effectuer un travail double également (Leff 2007). En outre, plusieurs phénomènes chimiques, dont le traitement théorique a su assimiler le concept d'entropie, comme la diffusion de deux gaz parfaits, sont typiquement décrits comme « irréversibles », impliquant une augmentation d'entropie sans toutefois indiquer une perte de travail (Planck 1908 : §134). Donc, si une interprétation valable n'est pas censée modifier la définition du concept, alors celle du travail est prise en défaut.

Ainsi, à première vue, l'entropie statistique aurait une portée, une extension plus grande comme concept que celui de « perte de travail ». Pourtant, il semble désirable de récupérer le principe de l'impossibilité du mouvement perpétuel de second type. Pour peu que le projet général de la mécanique statistique soit valable, un « arrimage » entre ces deux concepts devrait s'appuyer sur une certaine configuration de l'espace des phases, comme le fait justement le concept d'entropie statistique. Or, les concepts de chaleur et de travail en mécanique statistique ne sont pas aussi clairs qu'on le voudrait. En effet, une distinction de ces deux concepts semble exiger des considérations allant au-delà de la dynamique, comme la notion d'« observabilité » (Penrose 1970 : 14). Lorsqu'un système échange de l'énergie sous forme de travail mécanique, seulement un petit nombre de degrés de libertés est impliqué, de sorte que l'énergie transférée peut être mesurée par une observation directe des forces et des déplacements. Mais lorsque l'énergie est transférée sous forme de chaleur, plusieurs degrés de liberté inobservables sont impliqués, de sorte que l'énergie transférée doit être déterminée indirectement, comme avec la température. Penrose (1970 : 159) propose une distinction : lorsque les paramètres extérieurs déterminant l'évolution temporelle de l'hamiltonien peuvent être mesurés expérimentalement, alors ce changement dans le temps de l'hamiltonien, bref de l'énergie, peut être considéré comme une quantité de travail mécanique. Il en tire une conséquence : une condition nécessaire pour l'isolation thermique, dans le sens de la mécanique statistique, est que l'hamiltonien ne doit dépendre du temps que par ces paramètres extérieurs dont le changement dans le temps peut être considéré comme *connu* de l'observateur.

Ainsi, paradoxalement, la distinction entre ces deux concepts fondamentaux, du moins en thermodynamique, semble se trouver dans l'œil de l'observateur. Les capacités d'observation et d'intervention, à l'instar de celles des créatures surnaturelles comme le démon maxwellien, permettraient donc de saper cette distinction. Elles entraîneraient non seulement l'élimination de cette distinction mais aussi la possibilité d'un comportement réversible de l'entropie statistique. Il faut

toutefois faire preuve de vigilance ici. Ce ne sont pas, strictement, les capacités d'observation d'un être doué qui influencent le comportement des systèmes physiques. Les molécules ne collisionnent pas différemment que ce soit moi ou un « démon » qui les observons. Les capacités d'observation doivent être couplées à des capacités d'interventions, donc à des possibilités d'interactions causales tout ce qu'il y a de plus physique, pour que le comportement du système soit différent, mais seulement en ceci qu'il s'intègre dans un processus physique différent. Or l'observation est aussi un processus physique. Si les paramètres caractérisant un système sont connus, alors, trivialement, il est en principe possible d'intervenir sur ce système en modifiant l'un ou l'autre de ses paramètres. De fait, comme la quantité de chaleur et la quantité de travail dépendent du processus utilisé, un processus particulier peut produire plus ou moins de travail ou de chaleur selon le cas. Or ce processus ne peut être quelconque et dépend du système lui-même. C'est pourquoi les paramètres du système doivent être connus, une certaine quantité d'information est nécessaire, et plus cette information est précise, plus celle-ci est à grenage fin, plus l'intervention sur ce système peut être réglée de manière précise et ainsi « aplanir » l'écart entre les effets thermiques et mécaniques, et donc aussi l'irréversibilité. Cela signifie, plus précisément, qu'une intervention aussi précise, autant à grenage fin, se situe typiquement à l'échelle d'une molécule individuelle et joue donc sur le plan de la mécanique où il ne fait sens de discriminer les paramètres selon qu'ils peuvent être connus ou pas. En d'autres termes, le processus physique d'une telle intervention doit être décrit par des équations dynamiques, lesquelles font fi de la distinction chaleur-travail et de l'irréversibilité.

On constate ici à quel point il est facile de s'embourber dans ce genre de discussion entremêlant les concepts de travail, de chaleur, de probabilité et d'information. Il peut être tentant, par exemple, de conclure que les phénomènes irréversibles, pour lesquels la distinction chaleur-travail fait sens, ne sont le produit que de certaines limitations cognitives, qu'il ne s'agit que d'une question de point de vue. Pourtant il n'en est rien. Si le comportement d'un même système peut sembler différent à différents niveaux, à différentes échelles, c'est précisément en raison de différentes interactions (causales) avec ce système. Et l'information nécessaire à cette intervention n'est obtenue qu'au terme d'un processus physique décrivant aussi une interaction avec le système, laquelle est susceptible de modifier son état. Si le système est complexe au sens où plusieurs degrés de liberté sont nécessaires pour le décrire, alors il est évident qu'une plus grande quantité d'information, au sens large comme dans un sens plus technique, est nécessaire et pour le décrire et pour le contrôler. Un ensemble de particules se dirigeant toutes parallèlement dans la même direction peut être décrit exhaustivement avec peu de caractères, peu de bits, bref avec un court message contenant beaucoup d'information. Un tel système peut alors être considéré comme pouvant transférer son énergie « de manière cohérente » (Gould & Tobochnik 2010 : 18) et, pour une énergie donnée, il aura une faible entropie et une grande capacité à effectuer un

travail. C'est en ce sens quelque peu imprécis que l'entropie statistique peut décrire le « caractère aléatoire » (« *randomness* ») d'un système (Penrose 2004 : 690).

Ainsi, ce qui est premier ici n'est ni l'information, ni sa probabilité et encore moins un quelconque observateur, mais la disposition du système physique. Un certain observateur, pour lequel seulement certains paramètres macroscopiques peuvent être connus par une intervention à grenage grossier sur le système, n'est *d'abord* limité que parce qu'il a des limitations physiques et *ensuite*, c'est-à-dire conséquemment et corrélativement, parce qu'il n'a pas suffisamment d'information sur ce système. Le point crucial ici est qu'un éventuel observateur, soit un système capable de recueillir de l'information sur un grand système, doit justement en recueillir beaucoup afin d'ajuster ou de coordonner son intervention sur ce système de manière à aplanir la différence entre les effets du transfert d'énergie sous forme de chaleur et sous forme de travail. Une intervention à cette échelle peut être décrite uniquement par les équations de la mécanique où il n'y a pas de distinction entre travail et chaleur. Il faut alors remarquer deux choses.

D'abord, il est tout de même physiquement possible d'intervenir sur un système en l'absence d'information. Dans pareil cas, une intervention sans information préalable est peu susceptible de produire l'effet escompté d'une diminution d'entropie. Autrement dit, cet effet est peu *probable* au sens où plusieurs interventions ne mèneront pas à une diminution d'entropie. Par exemple, on peut imaginer une probabilité extrêmement faible pour que je passe ma main dans une boîte contenant un gaz et que ce faisant je frappe toutes les molécules de façon à les envoyer dans un coin précis de cette boîte. Ensuite, la discussion précédente montre qu'une observation, une mesure, n'est *pas* une procédure anodine à l'échelle microscopique car elle modifie le micro-état et qu'en plus elle modifie l'entropie globale du système et de son environnement ; comment, cela reste à voir toutefois, mais on retrouve les mêmes problèmes fondamentaux qu'en thermodynamique computationnelle (machine et Szilard et principe de Landauer) et en mécanique quantique (problème de la mesure).

Donc, les variables dynamiques caractérisant un grand ensemble de molécules sont les mêmes, que l'information qu'on en a soit faible ou élevée. Autrement dit, elles vont à la même vitesse qu'on le sache ou non. Ce n'est pas en ce sens que l'entropie est « subjective ». Mais une définition plus claire, une distinction plus serrée entre les concepts de chaleur et de travail est sans doute impossible sans faire intervenir le concept d'observabilité. Ces considérations sont donc très importantes pour notre discussion sur l'émergence (Chapitre 6). L'interprétation de l'entropie comme mesure ou capacité à effectuer un travail est très utile et très robuste au niveau macroscopique mais ne présente qu'une partie de l'équation. En somme, elle exige une circonscription claire du niveau (déjà un terme vague) dans la description du système à laquelle elle s'applique. Après tout, ce n'est pas surprenant, car un concept

n'est pas pertinent pour tout contexte ; il ne fait sens par exemple de parler de la température d'une molécule.

5.3.4.5 DISPERSION

L'interprétation de la « dispersion » ou de l'« étalement » – que l'on doit à ma connaissance à Leff (2007) – n'est pas étrangère, du moins dans son acception spatiale, à celle de « désordre ». Un gaz répandu dans tout le volume d'une boîte quelconque peut être dit « dispersé » et « désordonné » par rapport à un arrangement où toutes les molécules seraient rangées dans un coin de cette boîte, et a typiquement une entropie élevée. La métaphore est applicable, à première vue, aux différentes versions de l'entropie statistique. Dans le cas de l'entropie à grenage fin de Boltzmann ou de l'entropie de Gibbs, plus la fonction de distribution est étendue sur l'axe des abscisses, plus elle est dispersée en quelque sorte, plus l'entropie est élevée. Aussi, un micro-état où toutes les molécules seraient dans la même cellule, plutôt que d'être distribuées (dispersées) dans toutes les cellules, renvoie à une entropie très faible. Enfin, tel que mentionné, l'entropie de Shannon est maximale si tous les messages, tous les symboles ont la même probabilité, d'une certaine façon si les probabilités sont diffuses (dispersées) sur tous ces messages ou symboles.

Un point fort de cette interprétation est qu'elle place l'accent sur une *disposition physique* des systèmes avec la notion de « gradient », plutôt que sur de quelconques capacités cognitives. Les gradients, comme l'énergie, se présentent sous plusieurs variétés, que ce soit un gradient de température, de pression ou de potentiel chimique. Qui plus est, il semble bien y avoir une tendance dans la nature à l'élimination des gradients, ce qui peut être interprété comme une augmentation de la dispersion de certaines quantités physiques comme l'énergie. Deux corps mis en contact et à des températures différentes, par exemple, ont tendance à voir leur température s'uniformiser, la chaleur se dispersant uniformément dans les deux corps au fur et à mesure que le gradient de température disparaît. Cette tendance n'est toutefois pas complètement généralisable si l'on prend une acception large de ce que signifie un gradient. Il n'y a pas, par exemple, de tendance à l'élimination des gradients gravitationnels, bien que l'on peut se demander quelle propriété ces gradients soutiennent-ils. C'est-à-dire, si un gradient de température est l'expression d'une différence de température, alors un gradient gravitationnel exprime quoi au juste ?

Notons par ailleurs que cette tendance à l'élimination des gradients peut être freinée par l'imposition de contraintes, par des interventions. Un gradient de température peut ainsi être maintenu en plaçant un isolant thermique entre les deux corps ou carrément en ajoutant de la chaleur dans le système. Ce qui n'est pas sans rappeler les interventions permettant de faire diminuer l'entropie. Aussi,

la diminution d'un gradient de température par dissipation de la chaleur entraîne une diminution de la capacité d'effectuer un travail et une augmentation de l'entropie.

La métaphore de la dispersion permet aussi de recouvrir à première vue d'autres interprétations, déjà présentées. Ainsi, tel que mentionné, la tendance d'un gaz à se répandre dans le volume de son enceinte peut être interprétée comme une tendance à la dispersion. Un système dispersé ne présente pas de traits de symétrie et peut être vu comme désordonné. Et la quantité d'information pour le décrire nécessiterait un long message. Il est alors naturel de dire qu'une intervention sur un tel système pour le ramener à un état précis requiert beaucoup d'information et de précision.

5.3.5 Remarques conclusives

Devant la pléthore de définitions de l'entropie il n'est pas étonnant de constater que plusieurs interprétations ont été proposées. Déjà, les tentatives de la mécanique statistique sont aussi des interprétations à partir d'un autre langage, considéré fondamental et (plus ?) primitif (premier), celui de la mécanique classique. Les propositions sont donc nombreuses et on a vu qu'il n'y a pas de solution facile, pas de réponse à l'emporte-pièce.

Interpréter le concept d'entropie d'après le concept de probabilité découle du projet même de la mécanique statistique. Déjà Boltzmann considérait les « problèmes » de cette discipline comme étant des problèmes du calcul des probabilités (Chapitre 3). Mais devant le nombre impressionnant de discussions consacrées à l'interprétation de ce concept on est en droit d'être dubitatif quant au réel succès d'une telle interprétation comme pouvant favoriser la compréhension. Ce qui détermine les valeurs des paramètres thermodynamiques et macroscopiques détermine aussi le cadre formel dans lequel s'inscrit la gamme des micro-états possibles de la mécanique. Un état macroscopique peut donc être dans l'un ou l'autre de ces micro-états, ou encore il peut être dans telle ou telle région de l'espace des phases. Il semble donc naturel de définir ce rapport entre ce qui détermine l'état macroscopique et ce qui détermine la gamme des possibles microscopiques comme la probabilité de cet état macroscopique. De plus, compte tenu du *fait* que les systèmes thermodynamiques tendent à évoluer vers un état d'équilibre *et* qu'un tel état est typiquement décrit comme un état compatible avec la plus grande gamme de possibles, le plus grand nombre de micro-états possibles ou la plus grande région de l'espace des phases, et de loin, alors il semble aussi naturel d'expliquer cette tendance fortement corroborée comme un processus d'un état improbable à un état plus probable.

Mais cette interprétation est trompeuse. D'abord, on a vu que la méthode d'attribution de valeur de probabilités, qui définit aussi leur interprétation (critère de vérifiabilité), ne pouvait être réduite à l'un

des termes du schème objectif-subjectif, à moins de transposer l'observation de faits lors d'une mesure dans le champ de la psychologie du sujet ou d'ériger notre ignorance, par le biais du principe de raison insuffisante, au rang de principe formel justifié. Par conséquent, toute attribution de valeur de probabilité possède une part d'arbitraire, sans pour autant impliquer qu'aucune attribution de la sorte ne puisse être justifiée. Ainsi, la quantité Ω qui est associée à une équiprobabilité peut recevoir un appui empirique en ce que la plupart, pour ne pas dire la totalité des préparations mènent à une augmentation d'entropie, en conformité avec le principe entropique. Cependant, ce principe ne peut être convenablement appliqué si l'on fait abstraction et de la définition du systèmes (ses frontières) et de son isolation (ses interactions). D'éventuels arguments opérant de la sorte peuvent prétendre au bris de la connexion entre entropie et probabilité, mais en définitive c'est la validité du principe entropique qui est en jeu, puisqu'il serait constamment violé. Le concept de probabilité pouvant s'appliquer au concept d'entropie est celui de probabilité synchronique, où une mesure objective permet de définir des paramètres macroscopiques, mais aussi un espace des phases et un ensemble de micro-états compatibles avec cette mesure. En ce sens, ce qui est très probable est que le système se trouve dans un micro-état, dans des conditions initiales dont l'évolution déterministe selon les équations dynamiques garantie que l'entropie va augmenter – et cela se produit effectivement.

L'interprétation du désordre, pour sa part, trouve sa source dans certaines intuitions en ceci que l'ordre est davantage associé au travail qu'à la chaleur, mais aussi dans la théorie de la mécanique statistique au sein de laquelle on retrouve la définition microscopique d'un travail comme un transfert d'énergie « de manière cohérente » (Gould & Tobochnik 2010 : 18) ou impliquant « peu de degrés de liberté observables » (Penrose 1970 : 158). Il faut alors remarquer deux choses. D'une part, on constate que le niveau de description peut avoir une influence majeure sur la définition de certains concepts, car ceux de travail et de chaleur, pourtant robustes du point de vue de la thermodynamique, se voient associés à la notion d'observabilité et aux capacités cognitives de l'observateur en mécanique statistique. D'autre part, si l'ordre se définit par rapport à une règle, il faut savoir justifier une telle règle, qui est parfois conventionnelle, et comprendre ses implications. Même en justifiant cette règle, l'ordre et l'entropie ne covariant pas en raison de la non-extensivité de l'ordre, contrairement à l'entropie.

En ce sens, l'ordre d'un état peut être associé à sa probabilité, car déterminer l'ordre d'un état par rapport à une règle revient à *spécifier* cet état au sein d'une gamme d'états possibles et donc de rendre cet état spécial par rapport aux autres états de cette gamme. Par exemple, c'est bien en raison d'une règle conventionnelle que l'ordre alphabétique des livres d'une bibliothèque est établi. Parmi l'ensemble des états (des micro-états par rapport au tout qu'est la bibliothèque), ce dernier est spécial et il peut être mis en rapport avec l'ensemble des états possibles (soit la position des livres sur les étagères)

de manière à lui attribuer une probabilité. Afin de préserver cet ordre, des informations et des interventions sont nécessaires : il faut observer les livres et décoder ce qui les identifie (par ex. leur titre) et les déplacer au besoin. Cet exemple est illustratif mais tient même si les livres, contrairement aux molécules, ne se déplacent pas seuls. Il faut donc savoir bien appliquer le concept d'entropie (je rappelle l'exemple du jeu de cartes recevant de la chaleur et dont l'entropie thermodynamique augmente tandis que l'ordre des cartes demeure évidemment le même).

Le concept d'entropie peut donc être associé au concept d'information, mais ne peut s'y réduire. À moins de rendre cette association triviale en définissant ce dernier par le premier, ce qui est souvent le cas. Ce qui est important est de bien distinguer l'état du système de la connaissance que l'on peut en avoir, mais aussi des interventions et du contrôle sur ce système. Un système possédant une entropie élevée nécessite pour sa description un long message qui ne peut être compressé en raison de la dispersion de ces constituants (par ex. avec un message « toutes les molécules sont à droite »). Par conséquent, une intervention macroscopique (par ex. avec un piston) visant à obtenir un état spécifique (ordonné ou non) nécessite aussi beaucoup d'information car il faudrait intervenir (en supposant que cela soit possible) sur chacune des molécules (on peut imaginer un très petit piston). Mais qu'un quelconque sujet possède cette information ou non ne change rien à la disposition physique du système, comme celle lui permettant d'effectuer un travail. Il faut aussi comprendre, étant donné le mouvement incessant de ces molécules, à *a fortiori* si le système est ergodique, que n'importe quel micro-état est atteignable si l'on attend suffisamment longtemps. Pour aller plus vite, il faut de l'information *et* une intervention. En ce sens, dire qu'un système possède une entropie élevée, c'est aussi dire que la longueur du message nécessaire pour décrire le micro-état est grande – il s'agit de l'approche sémantique de l'entropie interprétée selon le concept de désordre. Cette approche occulte cependant le coût physique nécessaire à l'obtention de l'information. Dans l'approche substantielle de la connexion entre l'entropie et l'information, celle-ci est définie en termes d'entropie plutôt que l'inverse, où l'échange d'information est considéré du point de vue physique. Des résultats récents montrent ainsi que l'effacement de l'information est un processus dissipatif impliquant une augmentation d'entropie dans l'environnement. Ce qu'aurait sans doute permis de conclure une expérience de pensée faisant un usage cohérent du concept d'entropie. Encore faut-il éviter l'argument circulaire en tenant pour acquis la validité du principe entropique.

Les variables dynamiques caractérisant un grand ensemble de molécules sont les mêmes que l'information qu'on en a soit faible ou élevée. Ce n'est pas en ce sens que l'entropie est « subjective ». Une fois que les molécules sont dispersées, la capacité d'effectuer un travail est objective et l'information ne sera utile (pour obtenir un travail du système) que si elle couplée à une intervention, qui ne peut que modifier le système. La question revient alors à savoir comment traiter cette incertitude

découlant de la sous-détermination du micro-état suite à cette dispersion. Que ce micro-état existe et qu'il occupe un volume nul de l'espace des phases n'implique pas que toute caractérisation de la sous-détermination par une région de volume non-nul est nécessairement une question de préférence.

En somme, cette discussion quant à l'interprétation du concept d'entropie est riche d'enseignements, car chacune des propositions exposent un sens pertinent attribué à ce concept, tout en éclairant d'autres concepts aussi fondamentaux, comme la probabilité, le travail et l'information. Les difficultés de choisir une seule interprétation au moyen d'un concept encore plus primitif révèlent qu'il s'agit là d'un concept fondamental, mais qu'il faut encore l'explorer. Mais ce n'est pas une raison pour abandonner les efforts d'interprétation. « L'abandon raisonné, de concert, de termes d'usage métaphorique nécessiterait beaucoup d'efforts qu'aucun bénéfice immédiat ne justifierait. » (Morange 1999 : 527)

5.3.6 Stratégie explicative

Soit un système isolé ou adiabatiquement fermé consistant en un gaz confiné par une cloison dans la partie gauche d'un récipient. Si la cloison est retirée, le gaz va immédiatement se répandre dans l'entièreté du volume du récipient et le gaz s'est approché d'un état d'équilibre. Il a été vu (Chapitre 4) que la stratégie explicative de la thermodynamique consiste à caractériser un tel processus par l'augmentation principielle d'entropie qui atteint une valeur maximale à l'équilibre. Comme l'entropie d'un système isolé ne peut décroître, le système demeure dans cet état d'entropie maximale, à moins qu'il ne soit plus isolé. Ainsi, le second principe de la thermodynamique sous la forme du principe entropique rend compte de l'irréversibilité d'un tel processus en stipulant que l'entropie d'un système isolé ne peut décroître. On a vu aussi que l'objectif de la mécanique statistique est d'expliquer le comportement des systèmes thermodynamiques en dérivant des quantités statistiques par les lois dynamiques appliquées aux très nombreuses molécules constituant ces systèmes. Comme plusieurs l'ont noté (par ex. Khinchin 1949 ; Batterman 1998), il semble que son succès réside justement dans l'abstraction des détails microscopiques des systèmes qu'elle tente de décrire et d'expliquer. Bien que semblables à certains égards, il y a, comme pour le formalisme dont elles découlent, deux stratégies explicatives principales, celles de Boltzmann et de Gibbs, mais qui s'appliquent, de plus, à deux contextes différents, l'équilibre (entropie maximale) et le non-équilibre (augmentation d'entropie).

5.3.6.1 ÉQUILIBRE

La stratégie explicative d'un état d'équilibre en mécanique statistique est bien sûr conditionnée par l'approche adoptée, entre celle de Boltzmann et celle de Gibbs. Dans les deux cas il s'agit de refléter la phénoménologie de la thermodynamique en décrivant un état dans le cadre du formalisme de la mécanique classique, puis son changement plus ou moins monotone dans le temps, et enfin en montrant que cet état se maintient à une certaine valeur. La stratégie de Maxwell était d'identifier une distribution de probabilité des vitesses stationnaire censée représenter une condition d'équilibre d'un système isolé (Chapitre 3). Elle fut reprise par Boltzmann qui a tenté de montrer qu'une certaine quantité H , une fonction de la distribution de probabilité des vitesses, décroît dans le temps avant d'atteindre un minimum et s'y maintenir si la distribution est de type maxwellien. L'idée est donc que cette quantité représente l'inverse de l'entropie thermodynamique ($S_T \propto -H$), laquelle diminue avant d'atteindre un minimum à l'équilibre. Cette stratégie n'est donc valide que si le théorème H l'est d'abord (section 5.3.2.1).

La grande contribution de Gibbs est d'avoir permis une extension du traitement des systèmes à l'équilibre pouvant échanger matière et énergie avec leur environnement. Mais si l'on se limite à l'équilibre pour les systèmes isolés, donc pour l'ensemble microcanonique, alors on a dans le formalisme de Gibbs la stratégie explicative suivante (Sklar 1973 et 1993 ; Batterman 1998 ; Callender 1999). Il a été vu que la distribution de probabilité, ou densité, $\rho(x)$, joue un rôle majeur et qu'elle représente, à défaut de mieux à vrai dire, l'état du système, en d'autres mots l'explanandum³⁹. Comme précédemment, la représentation de cet état se fait au sein de l'espace des phases, plus spécifiquement de l'espace- γ , avec les coordonnées canoniques de position et de quantité de mouvement. Encore plus spécifiquement, comme le système a une valeur d'énergie fixée, l'espace des phases est confinée, réduite à une région bien précise qu'on appelle une « hypersurface d'énergie constante ». La distribution de probabilité s'applique donc à cet espace des phases réduit. Puisque l'équilibre est défini comme un état qui ne change pas dans le temps, qu'il est un invariant temporel, l'état mécanique statistique doit l'être aussi. Or, il n'y a aucun état microscopique, ou microcondition, ou un point représentatif, d'un système individuel qui présente cette propriété ; il suffit de penser au mouvement brownien, un mouvement « tumultueux » selon Planck. Il peut y avoir en revanche au moins une distribution de probabilité définie sur l'ensemble des états microscopiques présentant cette propriété. Un ensemble de Gibbs caractérisé de la sorte, avec une distribution de probabilité ayant certaines caractéristiques précises, est

³⁹ Il y a danger de confusion ici : dans le cas d'une explication réductive, l'« état » du système peut être à la fois l'explanandum du point de vue de la théorie réduite et l'explanans du point de vue de la théorie réductrice. Donc, à dire vrai, on veut expliquer l'état d'équilibre mais la densité de probabilité fait aussi partie ici de l'explanans.

alors en équilibre statistique. Il est ainsi possible de représenter une quantité physique, comme l'énergie, par une fonction en phase, $f(x)$, c'est-à-dire définie pour les coordonnées canoniques. À partir d'une distribution de probabilité $\rho(x)$ donnée, on peut calculer la moyenne de cette quantité physique, donnant ainsi la *moyenne en phase*, $\langle f \rangle$, de $f(x)$:

$$\langle f \rangle = \int f(x) \rho(x) dx.$$

De là plusieurs relations thermodynamiques à l'équilibre peuvent être obtenues.

Le choix ou la caractérisation de la distribution de probabilité devient donc primordial, ce serait même pour Penrose (1979 : 1940) le « problème fondamental de la mécanique statistique ». Dans la mesure où la stratégie explicative précédente permet d'inférer certaines valeurs pour certaines quantités physiques, ce choix peut faire l'objet d'un postulat. Ainsi, en construisant une distribution de probabilité, une densité déterminant justement comment les points représentatifs sont distribués dans l'espace des phases, la probabilité qu'elle désigne est généralement distribuée uniformément par rapport à la mesure de probabilité choisie, ce qui signifie qu'à chaque région de même volume dans l'espace des phases est attribuée une même probabilité. Il s'agit donc d'un postulat d'équiprobabilité. Bien que ce soit un postulat, ou plus précisément, bien qu'il soit considéré comme tel, plusieurs justifications visant à le rendre acceptable ou raisonnable ont été proposées. Comme le dit Sklar (1973 : 195), c'est là que certains physiciens, et les philosophes ne sont pas en reste, « ne parviennent pas à s'arrêter là » (alors qu'ils auraient peut-être dû). Par exemple, Tolman (1938) affirme que ce postulat est rendu nécessaire par l'incomplétude des connaissances quant aux états microscopiques. Cette justification n'est évidemment pas étrangère au principe de raison insuffisante, ou principe d'indifférence, une sorte d'argument de symétrie (déjà discuté ; voir aussi van Fraassen 1990 ; Sklar 1993 ; Uffink 1995).

La justification la plus célèbre, mais aussi la plus intrigante, est sans aucun doute l'*hypothèse ergodique* (discutée au Chapitre 3 et à la section 5.2.1). Avant de présenter cette hypothèse plus dans le détail, il importe de reconnaître ceci : bien que sa justification laisse à désirer et qu'il semble qu'elle soit inapplicable aux systèmes que la mécanique statistique cherche à décrire, elle résoudrait, supposant qu'elle soit valide, bon nombre de problèmes, dont celui de la justification du postulat d'équiprobabilité et ce, de manière complètement orthogonale au concept de probabilité. En substance, cette hypothèse stipule que la moyenne en phase d'une fonction censée représenter une quantité physique supposément observable est égale à sa moyenne temporelle. Une manière imagée de se la représenter consiste à imaginer une sphère parcourue par un point en mouvement, où le point est le point représentatif du système évoluant dans l'espace des phases, soit la sphère, et après un certain temps qu'on imagine suffisamment long, toute cette sphère aura été traversée par ce point en mouvement. En clair, cela

signifie que le système passe par tous les points de l'hypersurface d'énergie constante, par tous les états microscopiques compatibles avec son paramètre d'énergie.

Outre la justification d'une distribution et d'une mesure de probabilité particulières, la théorie ergodique permettrait aussi de justifier la procédure visant à faire la moyenne d'une fonction en phase. En effet, il faut justifier l'identification de la moyenne en phase avec le résultat d'une mesure expérimentale. Comme les mesures expérimentales de quantités physiques macroscopiques ne rapportent pas les valeurs instantanées des micro-états, ou même de ces quantités physiques, chaque mesure doit durer un certain temps qui peut paraître très court à l'échelle humaine, mais qui est très long à l'échelle microscopique où les micro-états subissent de nombreux changements, en raison par exemple des nombreuses collisions moléculaires. Par conséquent, on suppose que ce qui est expérimentalement accessible se limite à des *moyennes temporelles* sur de longues périodes. Ainsi, la mesure d'une quantité thermodynamique, macroscopique, est supposée durer plus longtemps que le temps requis pour la disparition des corrélations moléculaires (par exemple, les interactions à la collision), mais tout de même moins longtemps que le temps maximum au cours duquel la quantité macroscopique ne change pas. Or, ce temps maximum peut être considéré comme *infini* puisque l'équilibre est considéré comme un état invariant. De sorte que les moyennes temporelles f^* s'effectuent sur des temps infinis ($T \rightarrow \infty$) :

$$f^* = \lim [T \rightarrow \infty] (1/T) \int_0^T f(T_t x_0) dt,$$

où x_0 est un point en phase initial et T_t est une transformation mathématique, un opérateur d'évolution.

Il y a, bien entendu, quelques problèmes d'ordre physique (théorique) et philosophique avec l'introduction des temps infinis en mécanique statistique par le truchement de l'hypothèse ergodique (Jodoin 2011a). D'abord, la moyenne temporelle est intéressante parce qu'on suppose qu'elle est expérimentalement accessible, c'est-à-dire qu'on identifie la valeur d'une mesure empirique avec la valeur de la moyenne temporelle f^* de la fonction en phase $f(x)$. L'hypothèse ergodique veut que cette moyenne soit effectuée sur des temps infinis pour qu'elle puisse représenter la valeur d'équilibre de f . Cependant, la limite $T \rightarrow \infty$ ne nous dit rien à propos de ce qui se passe en un temps très long mais néanmoins fini, où cette expression peut dévier de beaucoup de la valeur limite à l'infini. Par conséquent, la théorie ergodique ne peut distinguer les systèmes à l'équilibre et ceux qui ne le sont pas. De plus, sur des temps infinis les systèmes peuvent posséder plusieurs états d'équilibre, comme dans le cas des transitions de phases, et donc il y aurait plusieurs valeurs invariantes pouvant représenter l'équilibre.

Ensuite, la moyenne temporelle f^* peut exister alors que le système n'est pas à l'équilibre. C'est-à-dire que la limite que représente f^* peut exister, alors que la limite de $f(T, x_0)$ n'existe pas. Par exemple, pour un mouvement périodique, la limite de la moyenne existe, en l'occurrence la moyenne des extrema pour un système simple, mais la limite de la fonction n'existe pas, justement en raison du mouvement périodique. Mais surtout, si un résultat de mesure d'un système dynamique correspond bien à la moyenne temporelle infinie, alors on voit mal comment il serait possible de mesurer et d'observer des changements chez ce même système. En effet, des mesures consécutives donnent souvent des résultats différents, autrement on ne pourrait observer des processus de non-équilibre. Cependant, la moyenne temporelle infinie reste la même pour un même système. On ne peut donc pas conclure que les mesures correspondent *toujours* aux moyennes temporelles infinies.

Enfin, empiriquement, l'équilibre du système physique est généralement atteint rapidement. Toutefois, le temps T nécessaire pour que f^* soit près de $\langle \hat{f} \rangle$ dans le cas microcanonique peut être énorme, de l'ordre de la « petite patience » de Poincaré, soit plus de 10^{100} années. Y a-t-il alors vraiment adéquation empirique ? L'infini en physique est un infini potentiel et non actuel : il devrait se limiter aux raisonnements par récurrence appuyant une généralisation ou une extrapolation, ou aux descriptions conditionnelles de la potentialité d'un système à poursuivre sans limite un comportement dans le temps et dans l'espace, comme avec le principe d'inertie. Si les idéalizations ne sont pas nulles et non avenues dans l'explication scientifique, elles doivent être utilisées avec précaution (Chapitre 2). Ici l'infini intervient plutôt pour satisfaire aux conditions de représentation mathématique de la propriété d'invariance temporelle.

Donc, la justification de l'identification des moyennes en phase sur des temps infinis et des quantités mesurées expérimentalement est « fallacieuse » (Sklar 1993 : 193). De plus, il n'y a pas de preuve pour l'applicabilité de l'ergodicité aux systèmes typiquement thermodynamiques, bref aux systèmes où la stratégie explicative est « requise » (Earman & Rédei 1996 : 70). Le problème est alors le suivant : pourquoi les hypothèses d'équiprobabilité fonctionnent si bien à donner les valeurs d'équilibre ? Plus particulièrement, il est décevant de ne pas avoir de justification pour l'équiprobabilité apriorique. On peut soutenir que cette justification réside dans le succès de la méthode, mais ce succès n'établit pas la vérité de la règle ni ne constitue en soi une explication. Ces difficultés suggèrent une position de retrait en qualifiant la tendance vers l'équilibre de « fait brut ».

5.3.6.2 AUGMENTATION D'ENTROPIE

Le *récit standard* (« *standard story* ») de l'*augmentation* d'entropie thermodynamique au sein de la mécanique statistique est le suivant. Les systèmes thermodynamiques tendent, lorsqu'ils sont laissés à eux-mêmes

un certain temps, vers un état d'équilibre et s'y maintiennent : il y a alors augmentation d'entropie vers l'équilibre, puis stagnation d'entropie une fois l'équilibre atteint. En mécanique statistique, l'état d'un système à l'équilibre est représenté par un état d'entropie élevé caractérisé par un grand nombre de micro-états compatibles avec les paramètres macroscopiques de cet état d'équilibre (grand Ω ou $\mu(\Gamma_M)$ ou ρ pour une région donnée). Or, on associe à l'entropie une *probabilité* en fonction du nombre de micro-états compatibles avec certains paramètres macroscopiques et mesurables, déterminés, de sorte qu'un macro-état caractérisé par un grand nombre de micro-états, donc par une valeur d'entropie élevée, est plus probable. Et, on l'a vu (section 5.3.4.2), un grand nombre de micro-états associé à un macro-état peut être interprété, encore que cela est discutable, comme un état de désordre. En conséquence, la tendance vers l'équilibre pourrait être expliquée par le principe voulant qu'un état aurait tendance à passer d'un état peu probable à un état probable, en d'autres termes d'un état ordonné à un état désordonné. Voyons de plus près ce que cela signifie.

Cette stratégie explicative de la mécanique statique s'articule autour d'un principe (équipartition probable), de deux postulats (statistique et dynamique) et d'une condition (d'évolution probable). Le *principe d'équipartition probable* ou *loi de Boltzmann* (Callender 1999 ; Uffink 2007 ; Frigg 2008 ; Brown, Myrvold & Uffink 2009) s'énonce comme suit : soit un instant arbitraire t et soit un système dont l'entropie de Boltzmann $S_B(t)$ est faible à cet instant, il est donc *très probable* qu'à un instant $t' > t$, $S_B(t') > S_B(t)$. Ce principe découle du principe entropique qui stipule que l'entropie augmente. C'est-à-dire que le principe d'équipartition probable tente de justifier ou d'expliquer ce principe entropique. Il appert donc qu'une justification de cette stratégie passe par une interprétation des probabilités. Plus particulièrement, la justification de l'équipartition pose de sérieux problèmes. Il semble insatisfaisant d'avoir recours uniquement au principe d'indifférence, ou à l'ergodicité, ou encore au principe d'entropie maximale.

Il est alors opportun de proposer une interprétation selon les termes exposés plus tôt (section 5.1.3), soit comme une sous-détermination synchronique dans le cadre d'une objectivation de certains résultats de mesures et de certaines descriptions nomologiques. Ici, le mot « certain » signifie que ces résultats de mesure et ces descriptions sont relatives non pas à un sujet (seulement de manière triviale⁴⁰) mais à un cadre théorique. En appelant à une probabilité synchronique toutefois, on ne règle pas le problème de l'augmentation d'entropie, celui de la condition d'évolution probable. Il a été vu qu'il est faux de dire que les systèmes passent d'états peu probables à des états plus probables et par conséquent cette assumption ne peut justifier l'augmentation d'entropie. Plutôt, un système thermodynamique isolé a

⁴⁰ Par exemple, pour éviter toute confusion, les lois newtoniennes sont considérées objectives même s'il a bien fallu un Newton, au demeurant un homme, un sujet, si génial soit-il, et même s'il en faut aussi un, même indirectement, pour les appliquer.

une forte probabilité synchronique d'être dans un micro-état évoluant de manière déterministe vers un micro-état associé à une entropie statistique élevée, associée à une probabilité synchronique élevée. Cela signifie qu'il n'y a pas de passage simple d'une probabilité synchronique à une probabilité diachronique. C'est pourquoi, comme le soutient Clark (2001), l'interprétation propensionniste est incompatible avec le déterminisme de la mécanique hamiltonienne gouvernant les molécules constituant les micro-systèmes de la mécanique statistique. L'interprétation fréquentiste pour sa part doit faire face à l'objection, déjà soulevée par von Mises (1931 : 519), voulant que l'évolution temporelle d'un système individuel ne constitue pas un collectif et en conséquence ne satisfait pas l'une des conditions imposées par cette interprétation (voir aussi van Lith 2001 ; Clark 2001). Enfin, il semble que d'autres tentatives particulières (par ex. les « mosaïques humiennes » de Lewis) ne sont guère convaincantes (Frigg 2008).

En dépit de ces problèmes d'interprétations du concept de probabilité, le principe d'équipartition probable exige à tout le moins un lien entre le formalisme défini par l'espace des phases et les équations dynamiques, d'une part, et les probabilités, d'autre part. Et cela est possible, du moins cela est proposé, au moyen du *postulat statistique*, ou quelque chose s'y rapprochant, qui va comme suit : soit M_t le macro-état du système à un temps t ; alors la probabilité à t que le micro-état du système soit dans une région B de l'espace des phases, incluse dans une région Γ_M composée de l'ensemble des régions associées à un macro-état tel que $B \subseteq \Gamma_M$, lesquelles sont définies sur l'hypersurface d'énergie Γ_E (où $B \subseteq \Gamma_M \subseteq \Gamma_E$), est $p(B) = \mu(B)/\mu(\Gamma_M)$. Associer de la sorte le volume en phase et la probabilité d'état peut sembler intéressant, voire même intuitif (Handfield (2011) nomme cette thèse la « théorie modale du volume [en phase] »). Mais ce qui est en définitive recherché est plutôt une justification de l'*évolution temporelle* des états, en particulier des états thermodynamiques dont l'observation nous révèle qu'ils tendent vers l'équilibre. Or ce n'est *pas* ce que fournit le postulat statistique, du moins dans sa forme actuelle. Ce qui n'est pas surprenant compte tenu du type de probabilité dont il s'agit, la probabilité synchronique.

Une solution, encore une fois, consiste à ajouter cette *condition* : les systèmes évoluent d'états probables vers des états plus probables. Et en vertu du postulat statistique, l'on est en droit de s'attendre, et même d'expliquer, qu'un système dont le micro-état possède un faible volume en phase et par conséquent une faible probabilité et une faible entropie évoluera vers un micro-état possédant un grand volume en phase et par conséquent une probabilité et une entropie élevées. Mais cette condition en cache deux autres (Brown, Myrvold & Uffink 2009) : (i) le flot dynamique T_t doit être tel que pour la vaste majorité (définie par rapport à la mesure μ) des micro-états ω_0 dans la région Γ_M à t_0 correspondant à la valeur initiale de basse entropie, le micro-état ω_t résultant de l'évolution de ω_0 soit associé à un macro-état $\Gamma_M(\omega_t)$ dont la mesure est $\mu(\Gamma_M(\omega_t)) \gg \mu(\Gamma_M(\omega_0))$ à t_0 ; et (ii) la mesure μ doit

avoir une signification probabiliste. Malheureusement, (i) requiert une justification qui demeure absente car elle provient du postulat statistique et (ii) est insuffisant (Lavis 2005). Comme le fait remarquer Planck (in Kuhn 1978) cité précédemment, le calcul des probabilités ne peut prédire l'évolution des états ni justifier un principe à l'effet que cette évolution procède d'états peu probables à des états plus probables.

En effet, cela contredit ce que l'on sait, et même ce qu'on a supposé à la base de la mécanique statistique, de la mécanique classique, c'est-à-dire de l'évolution déterministe des micro-états. D'abord, un micro-état évolue, justement de manière déterminée, vers un autre micro-état en raison des équations de mouvement, et non *pas* parce qu'il y aurait *plus* de micro-états ou un *plus gros volume* dans une certaine région de l'espace des phases. En clair, les micro-états, constitués d'un ensemble de molécules, sont gouvernés par les équations dynamiques ou ils ne le sont pas. Ensuite, en raison des équations dynamiques, et plus particulièrement de leur invariance par rapport au renversement du temps, la très grande majorité des trajectoires des micro-états passant par une région dont le volume en phase est faible, et par conséquent possédant une faible entropie, vont évoluer vers une région dont le volume en phase est grand vers le *futur*, mais elles vont aussi évoluer vers une région dont le volume en phase est grand vers le *passé*. En effet, si $\{x_1(t) \dots x_N(t)\}$ est une solution des équations dynamiques de Newton, alors $\{x_1(-t) \dots x_N(-t)\}$ l'est aussi. C'est l'objection qu'a fait Loschmidt à Boltzmann en 1876 (Chapitre 3).

Or, ceci vient contredire la plupart de nos observations. En effet, cela revient à dire qu'un système dans un état de faible entropie était très probablement dans un état de forte entropie dans un proche passé. Par exemple, selon cette ligne de pensée, un verre d'eau avec de la glace présentement était très probablement dans un état où la glace était plus fondue dans le passé. Cela mène évidemment, en général, à de mauvaises rétrodictions. Une solution possible et bien connue consiste à poser des conditions sur le passé de l'état actuel et donc à relier dynamiquement deux points de l'espace des phases : étant donné un état passé de plus faible entropie que l'état actuel, cet état va évoluer vers un état de plus grande entropie. Afin de prendre en compte tous les états potentiels, on peut poser une telle condition à l'état initial de l'univers, au Big Bang, qui serait selon cette hypothèse dans un état de faible entropie. Ainsi, *tous* les systèmes physiques à un moment donné auraient généralement une entropie plus faible qu'à un moment ultérieur. Einstein, Schrödinger et Feynman auraient adopté cette hypothèse (Callender 2008 ; plus récemment par Lebowitz 1993 ; Price 1996 ; Albert 2000 ; Goldstein 2001). C'est ce qu'on appelle l'*hypothèse du passé* (« *Past Hypothesis* »). Si le postulat statistique à propos du volume en phase est valide pour un état macroscopique M_p à t_0 , alors, pour tout temps $t > t_0$, la probabilité à t que le micro-état du système soit dans une région B de l'espace des phases, incluse dans

une région Γ_M composée de l'ensemble des régions associées à un macro-état tel que $B \subseteq \Gamma_M$, lesquelles sont définies sur l'hypersurface d'énergie Γ_E (où $B \subseteq \Gamma_M \subseteq \Gamma_E$), est

$$p_t(B|M_p \cap \phi_t(M_p)) = \mu(B \cap (M_p \cap \phi_t(M_p))) / \mu(M_p \cap \phi_t(M_p)),$$

où ϕ_t est le flot dynamique (du postulat dynamique) dans l'espace des phases. C'est donc une probabilité conditionnelle d'un événement A étant donné B telle que $p_t(A|B) = p(A \cap B)/p(B)$.

Il a plusieurs objections adressées à cette hypothèse. D'abord, il n'y a pas de consensus sur son statut, à savoir s'il s'agit d'une hypothèse sur des conditions initiales, auquel cas on se demande si elle doit être expliquée, ou s'il s'agit plutôt d'une loi ou d'un « fait brut ». Ensuite, Earman (1996) soutient qu'elle n'est « même pas fausse » (« *not even false* ») puisque les modèles cosmologiques actuels ne permettent pas d'attribuer clairement ni une probabilité ni une entropie à l'état initial de l'univers, et même si c'était le cas, ces mêmes modèles ne garantissent pas une augmentation monotone de l'entropie. Enfin, si cette hypothèse, supposant qu'elle soit valide, garantit une augmentation globale de l'entropie de l'univers, elle ne permet pas en revanche d'inférer une augmentation locale d'entropie, pour un sous-système de l'univers. En effet, une augmentation globale de l'entropie de l'univers est compatible *a priori* avec plusieurs diminutions locales. Autrement dit, que l'entropie totale de l'univers comme un tout augmente n'aide pas à expliquer la plupart des comportements d'un de ses sous-systèmes. Toutefois, il n'est pas évident que cette dernière objection soit compatible avec le postulat dynamique. Car un état de faible entropie, avec une forte probabilité synchronique, est un état qui évoluera de manière déterministe, selon un flot dynamique avec un hamiltonien indépendant du temps, vers un état de plus grande entropie. Cette objection doit alors justifier des sous-systèmes de l'univers, à son état initial, qui seraient dans des états particuliers, de faible probabilité, évoluant vers des états de plus faibles entropies, ou encore déterminer un flot dynamique avec un hamiltonien *dépendant* du temps, un hamiltonien rendant compte des interactions entre sous-systèmes. (Pour des discussions à propos de cette hypothèse, voir Price 1996 ; Leeds 2003 ; Parker 2005 ; Frigg 2008a et 2008b.)

Eckhart (2006) propose aussi une justification de l'augmentation d'entropie par la même stratégie de conditionnalisation des probabilités sur des événements passés. Il tente alors de fonder l'asymétrie causale sur l'asymétrie temporelle définie par l'augmentation à grenage grossier de l'entropie macroscopique de Boltzmann. Cette proposition ne place pas seulement ce concept d'entropie en très haute estime, ce qui est discutable sans être absurde, mais elle offre aussi une justification partielle de l'augmentation d'entropie en se basant sur des intuitions très fortes quant aux inférences qu'on tire généralement sur le passé et le futur. Ainsi, étant donné qu'un micro-état quelconque est, d'après les équations dynamiques, compatible avec une multitude d'états postérieurs mais aussi avec une multitude d'états antérieurs, les prédictions et les rétrodictions semblent impossibles. Or, on en fait tous les jours

avec un certain succès. De plus, on reconnaît assez bien qu'un effet peut avoir plusieurs causes mais qu'une cause particulière ne peut avoir qu'une gamme restreinte d'effets (la question du *nombre* de causes et d'effets étant toutefois, Eckhart le reconnaît, problématique). La raison se situerait dans l'augmentation à grenage grossier de l'entropie. Ainsi, une inférence rétrodictive du présent au passé serait inopérante étant donné la très grande gamme de causes possibles. Une inférence sur un événement passé doit procéder plutôt par une prédiction à partir d'un événement passé encore plus antérieur, ce qui revient à conditionnaliser les probabilités sur l'événement passé à expliquer (à l'instar de l'hypothèse du passé). Par exemple, un singe se déplaçant sur des échasses laisserait très probablement des traces sur le sable, mais il faudrait placer plusieurs conditions sur cet événement pour inférer, de simples traces dans le sable, qu'elles ont été effectivement produites par un singe se déplaçant sur des échasses. Eckhart en conclut que les inférences sur le passé comme celles sur le futur sont basées sur des prédictions et qu'en conséquence, puisqu'elles sont basées sur le même genre d'hypothèses et la même méthode d'inférence, le problème de l'induction n'est pas aussi problématique qu'on l'aurait d'abord cru. On constate donc que la justification de l'augmentation de l'entropie fait intervenir sans doute plus de questions que de réponse. Hormis ces problèmes, la suggestion de la considérer comme un « fait brut » ouvre, on le voit avec cette proposition, un nouveau champ d'analyse à de vieux problèmes philosophiques.

Albert (2000 : 65) critique aussi le postulat statistique, qui associe le volume en phase et la probabilité d'état, et donne l'exemple suivant : soit X la propriété d'être un appartement contenant une spatule et soit une grande collection d'appartements sur terre ; si le pourcentage de systèmes ayant la propriété X dont le micro-état réside dans une certaine sous-région de la région X de l'espace des phases est proportionnel au volume de cette sous-région, alors le ratio des pourcentages d'appartements dont la spatule est (par ex.) dans la salle de bain et de ceux dont elle est dans le tiroir de la cuisine devrait être égal au ratio de la quantité d'espace que ces deux contenants (soit la salle de bain et le tiroir de la cuisine) occupent. Mais ce n'est clairement pas le cas. Il est très rare (du moins sur terre) que la spatule soit dans la salle de bain. Ici toutefois, l'analogie n'est pas tout à fait justifiée. C'est-à-dire qu'on ne peut associer ainsi le volume en phase et le volume spatial comme le propose ce raisonnement analogique. En effet, il semble évident que l'espace seul ne représente pas l'état des objets comme la spatule. L'espace des phases, par contre, est censé représenter l'ensemble des états possibles des objets auxquels il réfère. Ainsi, un grand volume en phase correspond à un grand « nombre », au sens intuitif que généralise la notion de volume, d'états possibles. De plus, si le volume spatial pouvait représenter adéquatement l'état des objets comme la spatule, ou si d'autres variables pertinentes pouvaient caractériser ses états, il n'irait pas de soi non plus que l'espace des phases ainsi défini pourrait satisfaire les conditions d'un tel raisonnement analogique. En effet, il est raisonnable d'associer volume en phase

et probabilité que si la mesure de probabilité est bien définie. Encore une fois, en mécanique statistique, on considère généralement une mesure uniforme de sorte que les états possibles sont équiprobables. Dans pareils cas, l'association entre volume en phase et probabilité est justifiée. Mais il est évident que l'espace des phases caractérisé uniquement par la position spatiale, comme c'est le cas dans l'exemple de la spatule, ne peut recevoir une mesure de probabilité uniforme.

En somme, affirmer que les systèmes évoluent d'états probables vers des états *plus* probables est problématique. Plus spécifiquement, la loi de Boltzmann appuyée par le postulat statistique, lequel s'appuie sur la double condition (évolution probable et équipartition) présentée précédemment est problématique. Mais il n'est pas tant question de la « probabilité que la glace va fondre », mais plutôt qu'il est très probable que le verre d'eau contenant glace soit dans un micro-état garantissant de manière déterministe qu'il va fondre. Et cette probabilité est conditionnelle au macro-état présent. En clair, les systèmes n'évoluent pas d'un état peu probable à un état plus probable, ils évoluent plutôt d'un état théoriquement compatible avec un ensemble de micro-états indéterminés par la mesure vers un autre état aussi théoriquement compatible avec un ensemble de micro-états indéterminés par la mesure. Or, et c'est là le point crucial, la majorité des micro-états, ou si l'on veut du volume en phase, aussi considérés comme des conditions initiales, évoluent vers des états d'équilibre, d'où le plus grand volume dans l'espace des phases. Cela revient dire que, toute chose étant égale par ailleurs, il y a plus de chance qu'un système quelconque soit dans un état qui va évoluer vers un état d'équilibre, à condition qu'il demeure isolé.

Dans ce cas, il n'est pas possible (du moins si l'on veut être cohérent avec nos observations) de fournir, sur cette base, de rétrodiction valable. C'est-à-dire que, sur cette base toujours, on devrait rétrodire que les systèmes dans un état de faible entropie présentement étaient très probablement dans un état de forte entropie quelques instants avant. Et ce n'est clairement pas ce qui est observé. D'aucuns pourraient dire alors : « Au diable l'observation ! *All Hails Newton's equations !* ». Mais la mécanique newtonienne n'est pas aussi universelle qu'on voudrait bien le croire, car lorsque les masses en jeu sont très petites elle doit être remplacée par la mécanique quantique et lorsque les masses sont très grandes d'importants correctifs relativistes s'imposent. De plus, l'idéalisation des points matériels peut paraître inoffensive, mais il n'est pas absurde de penser que la configuration des molécules, qui est négligée, puisse être un facteur déterminant par l'effet combiné d'un très grand nombre de molécules. Ou encore, à l'inverse, que l'effet combiné d'un très grand nombre de molécules rende impertinents ces détails microscopiques de manière à ce que les régularités susceptibles d'être promues au rang de lois se trouvent à un niveau d'organisation supérieur ; autrement dit, il y aurait émergence. Dans les deux cas cependant, la mécanique statistique n'est pas d'un grand secours.

Pour conclure, l'idée n'est donc pas qu'un macro-état évolue d'un état peu probable à un état plus probable, mais plutôt qu'il soit très probable, synchroniquement, que le macro-état corresponde à un micro-état qui évoluera de manière déterministe vers un micro-état d'équilibre. Une intervention extérieure peut faire passer le système d'un état d'équilibre à un état de non-équilibre, mais presque invariablement le micro-état correspondant à cet état de non-équilibre évoluera (c'est-à-dire le calcul) vers un état d'équilibre correspondant au plus grand volume dans l'espace des phases. Cette stratégie explicative est donc de type réductif, mais partiel, en ce sens où il n'y a pas de fonction en phase présentant le caractère de non-décroissance et l'entropie plus élevée d'un état ultérieur n'est pas dérivée de la mécanique mais d'hypothèses probabilistes. Elle peut évidemment être couplée à une interprétation particulière de l'entropie, comme celle du désordre ou du manque d'information, où l'augmentation de l'entropie devient l'augmentation, respectivement, du désordre ou du manque d'information, avec les problèmes que cela comporte.

5.4 Autres disciplines

Il a été vu (Chapitre 3) que plusieurs disciplines scientifiques utilisaient ou récupéraient, de manière plus ou moins métaphorique, le concept d'entropie. Je vais me limiter ici à trois de ces disciplines, soit la théorie de l'information, la mécanique quantique et la biologie (physiologie et génétique des populations).

5.4.1 Théorie de l'information

Les théories de l'information visent à améliorer la transmission de messages ou de signaux dont le « contenu brut en information » est déterminé à partir de leur description suivant une chaîne de caractères donnés. Ce contenu, qui peut être celui d'un poème de Molière ou du génome d'un virus, ne détermine pas sa « valeur », laquelle est relative à un certain « but ». Différents buts donnent différentes théories de l'information et aucune théorie générale de l'information ne peut « prendre en compte tous les aspects pragmatiques qui déterminent la valeur d'une chaîne de caractères » (Delahaye 1999 : 531). Si l'on se fixe comme but de compresser au maximum la chaîne de caractères avec un certain mécanisme de calcul, on obtient la théorie algorithmique de l'information, et la théorie de l'information de Shannon en est la version probabiliste.

L'entropie discrète de Shannon $S_{s,d}$ est définie telle que

$$S_{s,d}(P) \equiv -\sum p_i \log p_i,$$

où $P = \{p_1, \dots, p_n\}$ est une distribution de probabilités, p_i est la probabilité d'un message m_i et 'log' est le logarithme en base 2 (Shannon 1949). L'entropie discrète de Shannon est *maximale* lorsque $P = p_1 = \dots = p_n = 1/n$, où n est le nombre de messages, et est nulle, donc *minimale*, lorsque tous les p_i sont nuls sauf une. Ainsi, tout changement vers une égalisation des p_i mène à une augmentation de $S_{s,d}$. Et si toutes les probabilités p_i sont égales, alors l'entropie discrète de Shannon équivaut à l'entropie combinatoire de Boltzmann à une constante près. L'entropie de Shannon peut être généralisée au cas continu, soit l'entropie continue de Shannon $S_{s,c}$, en optant pour une densité de probabilité $p(x)$ pour une variable x telle que

$$S_{s,c}(x) \equiv -\int p(x) \log p(x) dx.$$

Il y a une différence importante entre les entropies discrète et continue de Shannon. Dans le cas discret, la valeur de l'entropie est entièrement déterminée par la mesure de probabilité sur les messages, tandis que dans le cas continu, la valeur dépend des coordonnées utilisées pour décrire le message.

La théorie de Shannon lie donc le concept formel de l'information à celui de probabilité. La formalisation du concept d'information découle de l'imposition de certains critères intuitifs. D'abord, l'information peut être vue comme un processus impliquant un émetteur, un message et un récepteur. Ensuite, compte tenu d'une certaine distribution de probabilité P pour les différents messages, un récepteur quelconque devrait obtenir plus d'information d'un message peu probable que d'un message très probable, de sorte que ce gain en information devrait être une fonction décroissante de la probabilité p_i ; en d'autres mots, ce gain en information est plus grand, plus grande est la « surprise » (Balian 2003 : 15) du récepteur, si la probabilité du message est plus faible. Ainsi, une certaine fonction, permettant de mesurer la quantité d'information d'un message, devrait satisfaire le critère d'*additivité*, comme quoi la quantité d'information de deux communications indépendantes doit être la somme de l'information de ces communications; celui de la *monotonie*, où l'incertitude, soit l'inverse de l'information, croît si l'on augmente le nombre de messages; et celui de la *normalisation*, où la somme des probabilités est égale à l'unité, mais aussi, par convention, l'information moyenne obtenue de deux messages ayant la même probabilité représente une unité d'information, un « *bit* ». Il y a aussi d'autres critères ou axiomes (Klir 2006). La satisfaction de ces critères mène de manière univoque à l'entropie discrète de Shannon.

5.4.2 Mécanique quantique

L'entropie quantique (S_Q) est due à von Neumann (1955) :

$$S_Q = S_{VN} \equiv -k_B \operatorname{Tr}[\rho \ln \rho]$$

(voir aussi Chapitre 3). L'argument de l'entropie quantique à l'effet que l'entropie quantique S_{VN} correspond à $-\operatorname{Tr}[\rho \ln \rho]$ comporte deux étapes. Dans la première étape, il soutient que les mesures quantiques augmentent l'entropie thermodynamique du système mesuré. Dans la seconde étape, il cherche un équivalent mécanique quantique qui se comporte comme l'entropie thermodynamique et adopte S_{VN} . Selon Capek & Sheehan (2005 : 17), S_{VN} offrirait une « description plus fondamentale et fidèle de la nature » que l'entropie combinatoire de Boltzmann en ce qu'elle s'appuie sur une définition claire des micro-états quantiques contrairement au continuum de la mécanique classique. Personne ne remet en doute que la discrétisation de la mécanique quantique est plus adaptée au dénombrement des micro-états que tout ce que peut proposer la mécanique classique. Le problème est plutôt dans la signification de l'entropie quantique, laquelle est à chercher dans sa dérivation, sa preuve, loin d'être exempte de problèmes.

Cette preuve est généralement absente et la définition de S_Q est simplement postulée (Peres 2002 ; Maroney 2008). Celle que présente von Neumann (1955) repose sur une expérience de pensée (une autre). Plusieurs hypothèses sont nécessaires, mais sa principale, et donc son principal problème, est de considérer les états quantiques de la même façon que les gaz parfaits. Une boîte, dans lequel se trouvent des particules dans un état pur en ce qu'elles sont dans un état 0 ou 1 avec une probabilité déterminée à $1/2$, est divisé en deux par membranes semi-perméables, c'est-à-dire qu'elles laissent passer les particules selon qu'elles sont dans l'état 0 ou 1, et toutes les particules se trouvent dans une partie de cette boîte. Une mesure est d'abord effectuée sur chacune des particules sans enregistrement toutefois, laissant le système dans un état dit mixte où la moitié des particules est dans l'état 0 et l'autre moitié dans l'état 1. Les membranes semi-perméables sont alors utilisées pour séparer les états, de sorte qu'une partie de la boîte contient toutes les particules 0 et l'autre les particules 1, et l'ensemble des particules occupent alors tout le volume de la boîte. Cette opération est supposée être non-dissipative, sans dépense nette de travail et isentropique. Ensuite, le gaz est comprimé à son volume original de sorte qu'un investissement de travail (W) est nécessaire, dont la valeur est déterminée par $W = NkT \ln 2$ (section 3.4.4.3). Si l'on suppose que cette compression est isothermique et quasi-statique, cette quantité de travail investi dans le système est libérée sous forme de chaleur dans l'environnement, et l'entropie (thermodynamique) décroît de :

$$\Delta S = \Delta Q/T = -(NkT \ln 2)/T = NkT \ln 2.$$

Les particules sont alors retournées à leur état pur par une procédure supposément réversible. Et par hypothèse, comme tout état pur a la même entropie, celle-ci peut être fixée à zéro. Un cycle est prétendument complété.

La stratégie de von Neumann est basée sur un argument que Shenker (1999) appelle l'« argument arithmétique ». La première assomption de cet argument serait que l'entropie, dans cette expérience de pensée, ne change qu'à la compression et décroît d'une valeur de $NkT \ln 2$. La seconde assomption stipule l'entropie du gaz est la même à l'état initial et à l'état final puisqu'il s'agit d'un cycle. Par conséquent, l'entropie doit avoir augmenté quelque part. L'argument veut que ce soit lors de la mesure transformant le système d'un état pur à un état mixte. Cela implique qu'une mesure est un processus dissipatif, irréversible, où il y a augmentation de l'entropie *dans* le système. Car si cette augmentation se faisant dans l'environnement, il y aurait diminution locale sans compensation. (Ce point, c'est-à-dire où il doit y avoir augmentation d'entropie, est souvent négligé dans ce genre de discussion.) La seconde partie de l'argument consiste à choisir une quantité quantique qui covarie avec l'entropie dans l'expérience de pensée. Le choix s'arrête sur S_{VN} qui présente cette covariation mais aussi une forme analogue à l'entropie de Gibbs.

Shenker (1999) soutient que l'argument échoue en raison de son inconsistance puisque ses assumptions ne seraient pas appliquées uniformément à toutes les étapes de l'expérience de pensée : il y a une étape où le système est dans un état pur alors que l'argument ne lui attribue pas une entropie nulle, pourtant établie telle par hypothèse. En sorte que le bilan, à partir duquel on infère que la mesure implique une augmentation de l'entropie dans le système, n'autorise plus cette inférence. L'argument échouerait donc à montrer que S_{VN} est le corrélat mécanique-quantique de l'entropie. Comme ce contre-argument ne prouve pas la fausseté de la conclusion mais plutôt l'argument lui-même, Shenker soutient que *si* S_{VN} est bien l'entropie, ou « un genre d'entropie », alors le bilan prétendument corrigé implique une diminution globale d'entropie pour un cycle fermé, et donc la possibilité d'une machine à mouvement perpétuel. On en conclut que cette conclusion absurde rend les prémisses fausses, ces prémisses étant elles-mêmes la conclusion identifiant S_{VN} à l'entropie.

Henderson (2003) a répliqué à ce contre-argument, soutenant que celui-ci s'appuyait sur une mécompréhension de la notion quantique d'état pur. Ce contre-argument stipule qu'une mesure suivie d'une séparation des états quantiques par une procédure réversible avec des membranes semi-perméables implique que ces états peuvent être considérés comme purs et donc que l'entropie du système est nulle. Henderson réfute ce point. Le point de discordance concerne donc l'effet d'une mesure quantique. Typiquement, *avant* une mesure, un ensemble de probabilités est assigné à un système quantique de plusieurs particules pour les différents résultats de mesure possibles, et *après* la mesure, la fréquence relative de ces résultats *devrait* correspondre à ces probabilités et chacune des particules *devrait* être dans l'un ou l'autre de ces états. La description théorique de plusieurs mesures sur chacune des particules du système est celle d'une transition d'un état pur à un état mixte qui implique le passage à un état d'entropie plus élevée et ce, *avant* d'observer ou d'enregistrer les résultats de ces mesures. Au risque de provoquer des problèmes d'interprétation supplémentaires, cette situation pourrait être décrite de sorte que la mesure aurait comme effet d'« convertir ontologiquement » (Henderson 2003 : 294) l'état d'une particule dans un état particulier (0 ou 1). Il est difficile de vraiment faire sens de ce genre d'interprétation puisque le traitement théorique ne tient pas compte de ce « fait ontologique » (on me pardonnera cet abus de langage) à moins que le résultat ne soit observé ou enregistré. La théorie quantique se présente ici comme une théorie non pas sur la réalité au sens ontologique du terme mais comme une théorie de l'information qu'on peut en avoir. Ce qui se passe au niveau de l'entropie est alors pour le moins « paradoxal » (Lubkin 1987 : 524), car si la transition susmentionnée implique une augmentation d'entropie, le « fait » d'observer ou d'enregistrer le résultat de mesure implique une diminution d'entropie. Donc, ce qui devrait au départ être une discussion sur la signification du concept d'entropie se situe alors sur un terrain d'analyse beaucoup plus vaste mais néanmoins fondamental sur la signification d'une mesure.

Les considérations entropiques peuvent donc être une heuristique de choix à la compréhension et à l'interprétation d'une mesure quantique (Jodoin 2009 et 2010b). L'apparent paradoxe d'une modification soudaine de l'entropie soi-disant en raison d'une simple lecture des résultats de la mesure déjà effectuée se dissout si l'on objective cette mesure comme un processus physique, et pas seulement comme un « réajustement mental » (voir Caves, Fuchs & Sachs 2007). Mais les conséquences entropiques d'une mesure ne sont pas aussi claires qu'on voudrait bien le croire. Certains ont avancé l'idée que la mesure quantique produit une *augmentation* de l'entropie du système mesuré car celui-ci devient alors « couplé » à l'appareil de mesure qui possède plusieurs degrés de liberté (Jancel 1963 ; Partovi 1989). D'autres, à l'inverse, ont soutenu qu'il y a *diminution* de l'entropie car la mesure fournit de l'information et que ce gain d'information exige un rejet d'entropie quelque part (Lubkin 1987). Ou encore qu'une mesure implique un investissement de travail *dans* le système (Biedenharn & Solem 1995), ou plutôt que ce serait le système qui investirait un travail dans l'appareil de mesure (Shenker 1999).

5.4.3 Biologie

On a vu (Chapitre 3) que la récupération du concept d'entropie en biologie se présente sous deux aspects : (i) selon une approche *substantielle*, en ce sens que l'entropie thermodynamique est simplement appliquée à une classe de phénomènes particuliers, les êtres vivants, considérés comme des systèmes thermodynamiques (généralement) ouverts ; et (ii) selon une approche *analogique*, où une fonction d'un espace des phases (généralement) génotypique trouve des similitudes avec l'entropie statistique et s'insère conséquemment dans le formalisme de la mécanique statistique (voir les chapitres 6 et 7)⁴¹.

Comme exemple de l'approche substantielle, le phénomène du métabolisme animal est pertinent, bien que sa description se place parfois sur le plan métaphorique. Ainsi, l'« ordre » serait « caractéristique de la vie » (Reece, Urry & Cain 2005 : 145), ou encore la production d'entropie serait un « genre de mesure de l'ampleur de la violence des mouvements et des réactions se produisant dans la nature »⁴² (Aoki 1987 : 322). À titre d'exemple, ce phénomène chez un animal quelconque peut être décrit comme la production de molécules complexes et « ordonnées », comme l'amidon, à partir de la nourriture qu'ingère cet animal, tandis qu'il rejette d'autres molécules supposément « désordonnée », comme l'eau et le dioxyde de carbone, mais aussi de la chaleur, dans son environnement, dans un

⁴¹ Malheureusement, l'emploi, sur les modes analogique et substantiel, du concept d'entropie en biologie ne peut recevoir ici un traitement exhaustif pour des raisons d'espace.

⁴² « *kind of measure of the extent of violence of motions and reactions occurred in nature* ».

écosystème plus large. En clair, l'« ordre local » serait créé aux dépens du « désordre global ». Le processus physiologique de la photosynthèse est aussi pertinent, bien que la biophysique n'est pas entièrement expliquer le phénomène. Il consiste en la production de glucose et d'oxygène à partir du dioxyde carbonique de l'air et de l'eau du sol dans un environnement baigné de radiations solaires. D'abord, l'énergie des photons de la lumière solaire excite des électrons des molécules d'eau à un niveau d'énergie supérieur, séparant ainsi ces molécules en oxygène, rejeté dans l'atmosphère, et en hydrogène. Ensuite, l'hydrogène se combine au dioxyde de carbone pour former des molécules de glucose, ce qui diminue localement l'entropie et qui implique des réserves d'énergie de haute qualité pour les plantes. Le problème est que le premier principe exige de la chaleur pour la réalisation de ces processus chimiques tandis que le principe entropique demande un retrait de chaleur (Müller & Müller 2009 : 474). Afin de satisfaire le principe entropique, le processus doit être considéré dans un contexte plus large où la consommation d'eau et l'apport d'air sont prise en compte, ce qui mène globalement à une augmentation d'entropie.

De récents travaux ont ainsi tenté de définir une entropie en plaçant les variables de la génétique dans le formalisme de la physique statistique (Sella & Hirsch 2005 ; Barton & Coe 2009), exemplifiant ainsi l'approche analogique et reprenant une stratégie déjà mise en œuvre par Fisher (chapitre 3). Sella & Hirsch (2005) proposent ainsi un modèle mathématique de la dynamique évolutionniste à partir d'une analogie avec l'appareillage mathématique de la physique statistique. Sous certaines conditions d'ordre général, il serait possible de décrire une dynamique évolutionniste de type markovien menant à une distribution de probabilité stationnaire d'un génotype fixé, c'est-à-dire du génotype du plus récent ancêtre commun. Cette distribution de probabilité serait donc analogue à la distribution de probabilité de Boltzmann-Maxwell, en ce sens où elle décrit aussi une distribution stationnaire pour un système à l'équilibre. La dynamique évolutionniste, avant l'équilibre ou avant que le génotype de la population atteigne l'état stationnaire, est déterminée par une fonction d'énergie, G , que Sella et Hirsch nomme « fonction de la fitness libre » (« *free fitness function* ») :

$$G = \langle \ln(Wm) \rangle - 1/v \langle \ln(P) \rangle,$$

où $\langle \rangle$ représente l'espérance mathématique, W est la fitness moyenne, v est le coefficient de sélection et $\langle \ln(P) \rangle$ correspond à l'entropie S_H de la distribution de probabilité des génotypes. Cette fonction G s'accroît de manière monotone alors que la population évolue vers une distribution stationnaire, à l'instar du théorème H de Boltzmann où la fonction H (l'inverse de l'entropie) décroît de manière monotone : « La forme de cette fonction d'énergie est précisément celle de l'énergie libre [de Gibbs], et la maximisation de la fitness libre est précisément analogue à la seconde loi de la thermodynamique » (Sella & Hirsch 2005 : 9545).

Barton & Coe (2009) développent aussi une analogie entre physique statistique et dynamique génétique évolutionniste en la généralisant à une gamme plus vaste de modèles. En laissant tomber le critère d'une population dont les génotypes sont près de la fixation, les auteurs ajoutent à la fonction de fitness libre G une quantité U qui est une mesure de la diversité génétique ou de l'effet des mutations, tel que

$$G = \langle \ln(Wm) \rangle + \langle U \rangle + 1/\nu S_H.$$

Les trois termes de cette équation représentent respectivement les effets de la *sélection*, des *mutations* et de la *dérive* (« *drift* »). L'entropie S_H est fonction des fréquences des allèles présents à chaque locus pour une population se reproduisant sexuellement, et elle décrit la dispersion des micro-états, soit les fréquences des allèles. Les auteurs définissent aussi une entropie S_Ω qui est une fonction des variables macroscopiques, soit la moyenne du trait, la variance du trait et la diversité génétique (qui n'est pas une observable). Elle est proportionnelle, à l'instar de l'entropie de Boltzmann, au « logarithme du nombre de micro-états compatibles avec un état macroscopique donné » et lie par conséquent « les descriptions macroscopiques avec les descriptions microscopiques » (Barton et Coe 2009 : 319). Aussi, de manière analogue à l'entropie de Boltzmann, les systèmes (biologiques) évolueraient vers un état macroscopique qui est compatible avec plus de micro-états, soit vers une plus grande S_Ω : « les populations ont tendance à évoluer vers des états susceptibles d'être réalisés de plusieurs façons (soit vers des S_Ω élevées) » (Barton & Coe 2009 : 317). Or, il ne semble pas justifié d'inverser ainsi la stratégie explicative de la mécanique statistique, au demeurant déjà problématique, d'autant plus que la transposition du gain explicatif se base sur une analogie formelle. (Pour une discussion plus détaillée, voir Jodoin 2014a.)

5.5 Conclusion

Ni le second principe de la thermodynamique ni le principe entropique dans sa version thermodynamique n'ont à voir, bien entendu, avec la dynamique microscopique. Dans le premier cas, une distinction claire entre le travail et la chaleur est nécessaire, ce que ne permet pas la dynamique microscopique en général ou les équations newtoniennes ou hamiltoniennes en particulier. De sorte qu'une référence à un certain niveau d'observabilité ou de contrôlabilité, par rapport à des paramètres externes, est nécessaire afin de les définir distinctement. Dans le second cas, les concepts de chaleur et de température interviennent dans une définition basée sur un cas limite idéalisé où des quantités, soit le ratio de la quantité de chaleur transférée et de la température à laquelle s'opère le transfert, se compensent mutuellement, exemplifiant ainsi la réversibilité.

Les concepts statistiques de l'entropie opèrent sur un tout autre terrain, celui de la dynamique moléculaire inobservable. Ils doivent satisfaire deux propriétés fondamentales qui sont tirées de la formulation thermodynamique puis attribuées à sa formulation statistique, soit la non-décroissance, définie dans le temps pour les processus adiabatiques, et l'additivité, où l'entropie de deux systèmes est la somme de deux entropies individuelles. Ils sont aussi définis selon deux approches, celle de Boltzmann et celle de Gibbs. Dans le premier cas, l'entropie est propriété d'un système individuel définie à partir du théorème H selon lequel toute distribution évoluerait vers une distribution de type maxwellien puisque une quantité $H[f]$ décroît toujours jusqu'à atteindre un minimum stationnaire dont la distribution est justement de ce type. L'entropie continue de Boltzmann est définie à partir de cette quantité et simule le comportement de l'entropie augmentant dans le temps avant de demeurer constante à l'équilibre ; l'entropie combinatoire de Boltzmann procède à une discrétisation de l'espace des phases, tandis que l'entropie macroscopique de Boltzmann attribue un volume à cette discrétisation. L'entropie de Gibbs, pour sa part, procède par analogie formelle et adopte une stratégie faisant intervenir des ensembles infinis fictifs.

La conclusion de la discussion quant à l'interprétation du concept d'entropie peut sembler insatisfaisante, mais elle est en fait riche d'enseignements. En effet, et d'abord, cette discussion a révélé les difficultés d'interprétation qui se manifestent dans une multitude de propositions. Il a été vu que chacune d'entre elles expose un sens pertinent attribué à ce concept, elles contiennent en somme une « part de vérité ». Puisque le travail d'interprétation revient pour une large part à justifier le choix de concepts primitifs afin de mettre en lumière le concept interprété, cette multitude d'interprétations peut donc être considérée comme symptomatique du caractère fondamental du concept d'entropie (il suffit de penser aux difficultés d'interpréter le concept d'énergie, de proposer d'autres concepts plus

primitifs). Ensuite, les différentes justifications visant à asseoir la validité d'une interprétation particulière permettent en retour d'établir et de clarifier plusieurs liens qu'entretient ce concept avec d'autres tout aussi fondamentaux, comme ceux de probabilité et d'information. En prêtant attention aux subtilités de ces justifications on en vient à distinguer différentes approches (par ex. approches substantielle et sémantique) ainsi que différentes positions épistémologiques orientant l'interprétation. On a vu toutefois que l'interprétation de la dispersion était une candidate, sans être parfaite pour les raisons évoquées, évitant plusieurs des écueils typiques à ce genre d'entreprise et se rapprochant davantage de la représentation originale de la thermodynamique.

La stratégie explicative de l'entropie statistique est double en ce qu'elle présente des différences entre les cas d'équilibre et les cas d'augmentation d'entropie, qui nous ramènent aux fondements mêmes de la mécanique statistique. En tentant de définir une gamme de possibles en accord avec nos mesures empiriques, cette stratégie ne peut faire l'économie du concept de probabilité. Que celui-ci soit indispensable n'implique pas toutefois que sa mise en application soit exempte de problèmes. Car l'attribution empirique d'une valeur de probabilité ne peut recevoir la garantie d'une inférence à partir d'une fréquence relative ni d'un postulat de symétrie se ramenant au principe d'indifférence.

Dans le cas de l'équilibre, le raisonnement est plutôt simple et cherche un état microscopique, c'est-à-dire formulé à partir des variables de la mécanique, qui puisse légitimement être un substitut statistique qui serait temporellement invariant, puisqu'un état d'équilibre est défini temporellement invariant. On peut s'objecter à cette hypothèse en soutenant qu'après tout, on ne peut espérer que cette propriété pourtant évidente au niveau macroscopique ne l'est pas du tout au niveau microscopique et qu'en conséquence il faut changer de registre, modifier nos attentes. Justement, si le langage de la mécanique est différent de celui de la thermodynamique, alors pourquoi la propriété « invariant temporel » devrait être la même ? Mais comme il n'y a pas d'alternative, on l'adopte faute de mieux. Si l'on cherche une microcondition temporellement invariante d'un système individuel, alors on constate qu'elle n'existe pas, car elle subit d'incessantes fluctuations. Alors on peut se tourner vers une densité de probabilité définie sur un ensemble infini de sous-systèmes, qui possède cette propriété d'invariance temporelle. Le problème est alors de savoir laquelle, car il en existe en principe une infinité. La sélection de la distribution recherchée consiste en une distribution uniforme. Cette distribution uniforme maximise aussi la valeur d'une entropie préalablement définie. Maintenant, comment justifier ce choix ? On peut alors récupérer des arguments de symétrie avec le principe de raison insuffisante ou principe d'indifférence afin de conjuguer avec notre ignorance. On peut aussi établir un principe stipulant qu'il faut choisir la distribution maximisant la valeur de l'entropie. Enfin, on peut se tourner vers l'ergodicité, qui ne fonctionne que lorsque les temps sont infinis.

Dans le cas de l'augmentation d'entropie, aucune fonction de l'espace des phases ne peut répondre à cette exigence d'accroissement monotone dans le temps. La seule solution qui semble envisagée est la méthodologie du grenage grossier. Elle permet d'éviter le problème de l'invariance temporelle (on se rappelle ici que l'on souhaite que l'entropie change, augmente) découlant du théorème de Liouville, mais contredit la mécanique. Si cette méthodologie peut sembler arbitraire et subjective, elle reflète justement la « grossièreté » de nos interventions macroscopiques : on ne peut raisonnablement prétendre qu'une mesure de température est sensible à l'influence d'une molécule individuelle. Elle s'accorde de plus avec nos observations, *si* l'entropie combinatoire ou macroscopique est adoptée. Car, pour la vaste majorité des conditions initiales, un système isolé de son environnement évolue de manière déterministe vers un micro-état se trouvant dans une région de l'espace des phases exemplifiant une valeur maximale selon ces concepts d'entropie statistique. On peut alors objecter que ces observations sont impertinentes car, au départ, on cherche à les expliquer. On touche là à un problème de taille. Car on pourrait aussi soutenir que l'augmentation d'entropie tant recherchée aurait été introuvable, parmi l'infinité des possibilités de l'espace des phases, n'eût été de l'observation, de l'entropie thermodynamique.

L'idée n'est donc pas tant qu'un macro-état évolue d'un état peu probable à un état plus probable. Mais plutôt qu'il soit très probable, par une définition synchronique, que le macro-état corresponde à un micro-état qui évoluera de manière déterministe vers un micro-état d'équilibre. Il s'agit donc d'une explication réductive partielle. Une intervention extérieure peut faire passer le système d'un état d'équilibre à un état de non-équilibre, mais presque invariablement le micro-état correspondant à cet état de non-équilibre évoluera vers un état d'équilibre correspondant au plus grand volume dans l'espace des phases. En principe on peut calculer pour des systèmes relativement simples l'évolution du micro-état compatible avec une entropie élevée et en pratique on corrobore que l'état final d'équilibre correspond à un volume en phase élevé. Selon les définitions de l'entropie statistique, l'entropie est alors maximale. Reste à lui adosser une interprétation particulière selon le contexte, mais surtout en fonction des restrictions discutées précédemment. Cette stratégie explicative montre qu'une explication de type réductif est possible, mais il ne s'agit pas d'une réduction totale et l'attribution d'un statut d'émergence à l'entropie demeure donc concevable.

6 Entropie : explanandum ou explanans de l'émergence ?

RÉSUMÉ. L'émergence est parfois présentée comme ce qui expliquerait l'entropie alors que l'entropie expliquerait l'émergence. Tels quels ces deux énoncés ne peuvent être tous deux vrais. Face à la polysémie déconcertante des concepts d'émergence, d'entropie et d'explication, je soutiens que cet apparent paradoxe peut être résolu formellement selon le schème conceptuel suivant : l'émergence (en un sens *A*) explique (en un sens *B*) l'entropie (en un sens *C*) et l'entropie (en un sens *D*) explique (en un sens *E*) l'émergence (en un sens *F*). La solution revient donc à préciser *A*, *B*, *C*, *D*, *E*, et *F*.

Dans le cas de l'entropie comme *explanandum* de l'émergence (*A*, *B* et *C*), ce qui émerge est l'irréversibilité comme propriété essentielle de l'entropie thermodynamique, mais l'émergence ne peut être de type épistémologique (non-explicabilité). Mais si l'entropie n'est pas émergente et qu'une réduction est impossible (et il y a de bonnes raisons de le croire), alors on retombe sur un dualisme et la mécanique statistique devient inopérante. Donc l'émergence peut expliquer l'entropie thermodynamique selon (i) un sens fort, comme une modalité ontologique, (ii) en un sens intermédiaire selon l'émergence méthodologique, ou (iii) un sens faible, comme sa désignation comme membre d'une classe d'émergents.

Dans le cas de l'entropie comme *explanans* de l'émergence (*D*, *E*, et *F*), il faut distinguer l'approche *substantielle* de l'approche *analogique*. Dans le premier cas, l'entropie renvoie à une propriété macroscopique robuste et autonome pouvant être mobilisée au sein d'un explanans de l'émergence de nouvelles structures. Mais la stratégie explicative qu'elle soutient paraît inapplicable aux systèmes microscopiques. Dans le second cas, l'entropie exemplifie la réalisabilité multiple et elle est mobilisée au sein d'un explanans de l'émergence de propriétés à des niveaux supérieurs. Bien que la stratégie explicative que ce concept d'entropie veut porter semble incapable de justifier son augmentation dans le temps, la validité de cette stratégie analogique peut s'appuyer sur un solide formalisme, d'où l'importance de justifier l'analogie. La polysémie de ces concepts peut donc être fructueuse au sein de ce cadre explicatif de phénomènes divers et complexes, de la physique à la biologie.

I want to know am I the sky or a bird.

Tom Waits

Le progrès scientifique (ou ce que l'on juge être un progrès) procède autant de l'unité que de la diversité. De la diversité, avec la découverte de phénomènes nouveaux, et de l'unité par l'unification, au sein d'une classe ou d'un principe, d'un ou de plusieurs concepts représentant des phénomènes en apparence divers. Dans les deux cas la notion d'explication y joue un rôle important. Car cette unification est généralement le fruit d'une explication (par ex. l'électromagnétisme maxwellienne explique pourquoi la lumière est une onde électromagnétique parmi d'autres), mais il arrive aussi que ce qui semblait au départ le même phénomène dans des contextes différents sont en fait des

phénomènes distincts (par ex. on a longtemps confondu électricité statique et attraction magnétique¹). Le progrès peut aussi être le résultat d'un réarrangement de cette diversité avec la résolution d'une situation théorique intenable, voire paradoxale, où des positions divergentes en apparence irréductibles se disputent le même rôle explicatif, comme dans le cas de paradigmes concurrents.

Tel qu'indiqué plus tôt, la motivation initiale de ce projet est ce constat – apparemment paradoxal – à l'effet que l'émergence expliquerait l'entropie mais aussi que l'entropie expliquerait l'émergence. Tels quels ces deux énoncés ne peuvent être tous deux vrais. Autrement dit, ce qui explique semble aussi être expliqué de sorte qu'un concept, adjoint à des hypothèses particulières au sein d'un explanans, s'expliquerait lui-même. Le travail d'analyse et de désambiguïsation de cet apparent paradoxe doit alors faire face à la polysémie rampante des notions d'émergence, d'entropie et d'explication. Chacune de ces notions possède en effet sa propre taxonomie (Chapitres 2, 4 et 5). Nul doute donc qu'elles aient entraîné un lot d'assertions et de thèses les plus diverses et souvent contradictoires. La multitude de significations et d'interprétations de chacune d'entre elles prises individuellement déchaîne une multiplicité intenable de sens pouvant être accordés à leur combinaison au sein des deux énoncés précédents. Un paradoxe étant un énoncé sans solution, le présent paradoxe se dissout devant la solution consistant à assigner la bonne signification au bon terme. Il faut alors faire preuve de prudence et apporter les nuances et les précisions qui s'imposent.

La solution formelle est donc la suivante : l'émergence (en un sens *A*) explique (en un sens *B*) l'entropie (en un sens *C*) et l'entropie (en un sens *D*) explique (en un sens *E*) l'émergence (en un sens *F*). Autrement dit, l'émergence peut (dans certaines conditions *X*) expliquer l'entropie, et l'entropie peut (dans certaines conditions *Y*) expliquer l'émergence. Je propose alors de préciser adéquatement *A*, *B*, *C*, *D*, *E*, *F*, *X*, et *Y*.

En ce concerne l'entropie comme une *explanandum* de l'émergence, le débat s'articule autour de la tentative de traduire le concept d'entropie thermodynamique au sein de la mécanique classique, qui, étant donné des contraintes épistémiques relativement évidentes, s'adjoint le concours des statistiques et des probabilités. Il s'agit donc d'une entreprise réductionniste et de son succès dépend largement l'exemplification du concept d'émergence (Chapitre 2). Évidemment, le succès ou l'échec de moyens épistémiques ne constitue pas un frein à diverses affirmations métaphysiques et le concept d'émergence n'y fait pas exception. Il s'agit plutôt de justifier le cas échéant ce genre d'affirmations. Mais il faut aussi établir les conditions de succès d'une réduction et ce que cela implique pour le concept d'émergence. Or, affirmer un cas d'émergence revient déjà à caractériser une situation théorique devant ces efforts

¹ Ces phénomènes ont par la suite été réunifiés, mais pas uniquement sur la seule base de l'observation d'une attraction et a plutôt exigé tout un arsenal théorique avant de montrer que la même force était à l'œuvre.

réductionnistes. Ce succès et ces efforts sont jugés à l'aune d'une théorie de l'explication car généralement une réduction implique une explication.

En ce concerne l'entropie comme un *explanans* de l'émergence, deux approches sont ici distinguées, soit l'approche substantielle et l'approche analogique. Le premier cas renvoie à la stratégie généralisée du principe entropique (Chapitre 4), où, en termes simples, un système consomme de l'énergie et rejette de l'entropie afin de former une sorte d'organisation structurale. Devant le constat de cette organisation (qu'on appelle aussi « ordre », « complexité », etc.), comme ce qui est observé dans le domaine du vivant, plusieurs y ont vu une violation du principe entropique interprété comme un accroissement du désordre. Au contraire, le phénomène de la vie est désormais considéré comme un allié important du principe entropique et du second principe de la thermodynamique, puisque les composés chimiques biotiques génèrent beaucoup plus d'entropie qu'ils ne le feraient dans un état abiotique (Capek & Sheehan 2005 : 308). Une interprétation et une application rigoureuse du principe entropique permettent donc de voir qu'il n'y a pas violation.

Cependant, la physique – ni la biologie d'ailleurs – n'offre pas d'explication de cette tendance assez bien répandue vers un accroissement de l'organisation et de la complexité, des cellules vivantes aux systèmes galactiques. Si la généralisation du principe entropique, en supposant qu'elle puisse correctement répondre aux défis auxquels elle fait face dans les situations de non-équilibre, permet de conserver ce principe et même de l'appliquer à des phénomènes divers, et donc qu'elle *autorise* ce genre d'accroissement de l'organisation, en revanche elle ne l'*explique* pas. Qu'un système soit ouvert est une raison suffisante pour affirmer la possibilité nomique d'une diminution locale d'entropie, pas pour expliquer l'émergence de l'« ordre », d'une « structure ordonnée », d'une « organisation complexe », ou quelque chose du genre. C'est pourquoi certains ont proposé une « quatrième loi de la thermodynamique » afin de rendre compte de l'accroissement de l'organisation et de la complexité. Kauffman (1993, 1995 et 2000 ; Kauffman *et al.* 2007), par exemple, adopte cette approche dans son programme visant à expliquer le comportement des systèmes adaptatifs. Selon ce dernier, la complexité des systèmes biologiques résulterait d'une auto-organisation et d'une dynamique loin de l'équilibre, donc à faible entropie, en complémentarité avec la sélection naturelle. Il suggère ainsi une « nouvelle mécanique statistique » (1993 : 182) où l'espace des phases incorporerait des éléments couplés, en interaction. D'autres (par ex. Stewart 2003) affirment que dans pareils cas le second principe ne s'applique tout simplement pas.

Le concept d'entropie est donc mobilisé au sein de stratégies explicatives diverses mais néanmoins fort ambitieuses en ceci qu'elles traversent généralement plusieurs disciplines. Un signe d'une approche multi-, inter- ou transdisciplinaire. Examinons d'abord la complexité, où elle s'y applique souvent.

6.1 Systèmes complexes

Aucune définition de la complexité ne fait aujourd'hui consensus. Dans certains contextes elle réfère à ce qui est « difficile » (plus proche de « compliqué »), dans d'autres elle signifie le contraire de « simple ». La principale raison de cette absence de consensus est sans doute qu'il s'agit d'une notion primitive qui ne peut donc être codifiée aux moyens d'une définition explicite en termes de concepts préalablement signifiants et compris (Agazzi 2002 ; Chapitre 2). C'est pourquoi une définition précise, en fonction de conditions nécessaires et suffisantes, est évitée. Selon la conception classique de Simon (1962), un système complexe est un système constitué d'un grand nombre d'éléments interagissant de manière non triviale ou non simple. On y associe des notions telles que la non-linéarité, la non-additivité, l'irréversibilité, la brisure de symétrie, l'ordre, l'émergence, la hiérarchie, l'historicité, etc. (Hooker 2011). La raison pour s'intéresser aux systèmes complexes est double. Leur étude s'est développée en un champ de recherche à part entière (Edmonds 1999 ; Standish 2001 ; Bishop 2011) et on les considère tout à tour comme une condition à l'émergence ou comme l'une de ses conséquences. Ils sont peut-être en voie de devenir la norme plutôt que l'exception dans la pratique scientifique et d'amener une révolution de l'ordre de celle qu'a connu la physique il y a un siècle (Hooker 2011a).

Évitant, justement, une définition en fonction de conditions nécessaires et suffisantes, Rickles (2011) identifie, pour les systèmes complexes, un triplet de caractéristiques² : (i) ils doivent contenir un certain nombre (indéfini) de sous-unités (sous-systèmes), (ii) celles-ci doivent être interdépendantes (du moins pour un certain temps), et (iii) l'interaction entre sous-unités doit être non-linéaire. Les propriétés d'un système complexe (unité ou tout) sont considérées survenir sur celles des sous-unités (parties) qui le constituent. Mais lorsque ces propriétés sont « nouvelles », si elles transcendent en quelque sorte celles des sous-unités, alors elles sont dites « émergentes » (Chapitre 2). Il y aurait alors, dans certains cas (auto-organisation), une causalité descendante (Chapitre 2) en sorte que les propriétés (leurs instances) des sous-unités sont modifiées selon le comportement du système complexe qu'elles déterminent conjointement. Voici donc quelques caractéristiques (dont aucune ne fait vraiment consensus) des systèmes complexes.

Hiérarchie. Les systèmes complexes sont typiquement composés de plusieurs sous-systèmes (ou sous-unités) présentant plusieurs degrés de liberté. Ils présentent alors ce qui a été discuté plus tôt (Chapitre 2), soit des niveaux d'organisation permettant une description d'un ensemble de propriétés et de régularités suffisamment stable pour être déterminées de manière robuste. Cette stabilité n'est pas

² Rickles « hésite à les appeler des conditions nécessaires ».

absolue, mais implique que de petites perturbations ne modifient pas le comportement global du système et aussi, souvent, une certaine adaptabilité du système devant les modifications de son environnement (parler de l'environnement du système implique déjà une hiérarchie de niveaux). Souvent, la possibilité de distinguer différents niveaux et de former une hiérarchie découle de la présence d'un *processus irréversible* (Chapitre 4). Cependant, les multiples interactions non-linéaires entre niveaux d'organisation obscurcissent les distinctions, pourtant utiles et parfois indispensables, entre parties et tout, constituants et niveaux, système et environnement, etc. (Bishop 2011).

Non-linéarité et non-additivité. Une équation est additive si une combinaison numérique de ses solutions est aussi une solution. Le problème de sa résolution peut donc être ramené en une collection de problèmes plus simples. Dans ce cas, aucun comportement dynamique complexe n'est possible car la non-additivité est une condition nécessaire à ce genre de comportement (Hooker 2011). La non-additivité caractérise des équations non-linéaires, lesquelles représentent des interactions non-linéaires par rapport à un paramètre donné si la force d'interaction ne varie pas proportionnellement avec ce paramètre. La nouveauté, pour ainsi dire, résultant de la non-additivité fait dire à certains que la non-linéarité est « source de phénomènes émergents » (Mainzer 2005 : 159).

Prédictibilité. Le comportement des systèmes complexes est souvent difficilement prévisibles, principalement en raison de leur sensibilité aux conditions initiales, qui sont d'ailleurs difficilement déterminables en raison du grand nombre de degrés de liberté qu'il présentent, et en raison aussi des multiples interactions avec leur environnement. En général, la non-additivité entraîne un accroissement rapide des incertitudes dans le temps.

Probabilité. Les probabilités jouent désormais un rôle majeur en sciences et l'étude des systèmes complexes ne fait pas exception. Elles permettent ainsi de « rendre compte du fonctionnement intrinsèques des systèmes complexes, de dévoiler leurs régularités et les aspects de leurs structures »³ (Rickles 2011 : 536). En effet, les effets non-linéaires posent de sérieuses contraintes à la prédictibilité des systèmes qui les manifestent et des prédictions probabilistes sont d'un grand secours (Bishop 2011). On peut par exemple former un ensemble de conditions initiales à partir de l'état supposé du système et simuler son évolution dans le temps avec certaines mesures d'incertitude. Il s'agit d'une stratégie largement employée en mécanique statistique.

Ordre et organisation. La notion d'ordre possède plusieurs significations parfois difficiles à démêler. Elle renvoie souvent à un aspect subjectif car elle réfère à la capacité d'un observateur à déceler une configuration familière ou des traits de symétrie (Chapitre 5). Généralement, une plus grande symétrie

³ « to ascertain the inner workings of complex systems, to uncover their regularities and aspects of their structure ».

implique un plus grand ordre. On peut distinguer la notion d'ordre de celle d'organisation en ceci que la première désigne un haut degré d'uniformité interne (par ex. un cristal), tandis que l'organisation désigne « l'inter-articulation des composants d'un système » (Hooker 2011 : 29) ou la « collection de relations parmi les composants d'un système » (Bunge 2003 : 19). La définition technique de l'ordre est souvent celle de la théorie algorithmique de la complexité : l'ordre (« *orderedness* ») d'un motif est l'inverse de la longueur de sa plus courte description, ou la plus compressée. Autrement dit, un motif ou une configuration est ordonnée si sa description requiert peu d'information (bits). En ce sens, un gaz est désordonné puisqu'il requiert pour sa description microscopique de l'information sur chacune de ses molécules. En revanche, un laser serait plus ordonné étant donné la cohérence des photons, lesquels montrent ce qu'on appelle une grande pureté spectrale. Mais les molécules d'un gaz et les photons d'un laser ne présentent pas un haut degré d'organisation, comme dans le cas d'une cellule vivante ou d'un moteur d'avion. Enfin, le degré d'uniformité interne peut être évalué par rapport à une règle définissant cette uniformité (par ex. des symétries), mais le non-respect de cette règle n'implique pas que le système ou le motif soit chaotique ou désorganisé. Tel que discuté (Chapitre 2), Denbigh (1989a : 328) distingue ainsi le concept d'ordre où il y a un accord complet à une règle et le concept d'ordonnement (« *orderliness* ») comme mesure de l'écart d'avec cette règle. Mais ni le degré d'ordre ni la profondeur organisationnelle (mesurée par la collection d'inter-relations de sous-unités) ne capturent vraiment la notion de complexité (Hooker 2011).

Auto-organisation. L'auto-organisation est souvent définie, en termes généraux, comme l'apparition (émergence) de structures complexes, fonctionnelles, résultant d'interactions entre un système et son environnement (Hoelzer, Smith & Pepper 2006). Ou encore, une condition à l'auto-organisation serait que le système subisse une assez grande *bifurcation* (Hooker 2011a), soit une instabilité structurale menant à un changement de sa dynamique globale, une modification de son « paysage attractif » (régions d'équilibre de son espace des phases). Une bifurcation est généralement le résultat d'une *brisure de symétrie*, soit une non-invariance selon une opération de transformation (comme un renversement spatial ou temporel). Les cellules de Rayleigh-Bénard (section 6.3.1.1) en sont un exemple. Selon Hooker (2011a : 26), le résultat de cette brisure de symétrie serait une condition à « l'émergence de l'ordre et de la complexité »⁴. Mais, selon la définition précédente de l'ordre, cela est vrai par définition, car l'ordre serait une possibilité spécifique, définie par une règle (celle de la complexité) parmi une gamme de possibles. Malaterre (2010) distingue l'auto-organisation « structurelle », relative

⁴ À titre d'exemple illustrant l'intrication de ces concepts : « *Phase transitions in dissipative systems generate a bifurcation tree with emerging structures of increasing complexity. In this context, the degrees of increasing complexity are defined by the increasing bifurcations that lead to chaos as the most complex and fractal scenario. Each bifurcation illustrates a possible branch of solution for the nonlinear equation. Physically, they denote phase transitions from a state of equilibrium to new possible states of equilibria. If equilibrium is understood as a state of symmetry, then phase transition means symmetry breaking being caused by fluctuational forces.* » (Mainzer 2005 : 157).

à l'apparition de structures spatiales spécifiques, et l'auto-organisation « fonctionnelle », relative à la fonctionnalité de certaines structures vivantes, comme le métabolisme. Bien qu'utile en biologie, cette distinction doit être accompagnée de certaines nuances, car, d'une part, plusieurs structures différentes peuvent mener à des fonctions différentes, et, d'autre part, il est possible de subsumer des fonctions différentes sous un même concept, comme « aider à survivre » ou encore, comme on le verra plus loin, « faciliter les flux physiques ». Enfin, pour Bennett (2003 : 35), les processus irréversibles présenteraient une propension à l'auto-organisation, à la « génération spontanée de complexité ». On comprend donc que l'auto-organisation implique la complexité, bien que la définition de cette dernière demeure imprécise.

On constate que ces caractéristiques, dont la liste n'est pas exhaustive, se recoupent ou s'impliquent mutuellement. On constate également qu'elles dressent de sérieuses difficultés à certaines conceptions scientifiques et philosophiques classiques. L'une des conséquences importantes de ces constats est la nécessité d'impliquer plusieurs disciplines dans certaines stratégies explicatives (interdisciplinarité) ou de former de nouveaux concepts de manière conjointe et complémentaire (transdisciplinarité).

6.2 L'entropie comme explanandum⁵

Dans les différentes versions statistiques de l'entropie, il n'y en a aucune qui présente le caractère de la non-décroissance exemplifiant l'irréversibilité, essentielle au concept d'entropie thermodynamique, à moins d'y adosser des hypothèses fort discutables. Retrouver cette propriété exige en effet des stratégies et des concepts souvent au statut incertain et qui sont orthogonaux à la mécanique classique censée fonder ces versions que l'on pourrait qualifier de microscopique (Chapitre 5). Ainsi, dans la méthodologie du grenage grossier, bien que la partition de l'espace des phases ne soit pas soumise au théorème de Liouville, et par conséquent elle n'est pas soumise à l'invariance temporelle, celle-ci est souvent considérée comme arbitraire ou subjective puisqu'elle dépend des capacités observationnelles de l'observateur. De plus, on peut soutenir que la procédure même du grenage grossier viole les contraintes dynamiques du système et que par conséquent, à l'instar de la méthodologie des ensembles, on ne rend pas compte des macro-états thermodynamiques en fonction des micro-états dynamiques. Si l'on endosse l'hypothèse que l'entropie statistique est une fonction des variables dynamiques d'un système individuel mais aussi une fonction du temps ayant une valeur définie à chaque instant, en raison des équations de la dynamique hamiltonienne l'entropie ne peut pas croître de manière monotone, donc elle ne peut pas satisfaire le critère de non-décroissance.

Il s'agit, en somme, de bonnes raisons de croire qu'on ne puisse offrir une explication réductive de l'entropie thermodynamique et donc d'affirmer qu'il s'agit d'un cas d'émergence (Chapitre 2). Comme cet échec est celui de moyens épistémiques, cette émergence serait de type épistémologique. Cependant, l'échec de la réduction semble éliminer d'emblée la prétention de la mécanique statistique d'expliquer les phénomènes thermodynamiques. En effet, si une réduction peut avoir un pouvoir explicatif, prétendre que le niveau microscopique explique le niveau macroscopique *par* l'émergence entendue comme «absence de réduction» est dérisoire. On a vu d'ailleurs que l'émergence épistémologique est parfois définie comme l'explicabilité de ce qui émerge par ce qui est basique. L'explicabilité permettrait ainsi de caractériser la notion de nouveauté ou encore l'effet de surprise dont le concept d'émergence est censé rendre compte. En ce sens, l'entropie serait peut-être un cas particulier d'émergence, mais l'émergence n'expliquerait pas l'entropie. Il demeure possible toutefois que le concept d'entropie, ou plus particulièrement l'irréversibilité thermodynamique, puisse être expliqué par autre chose qu'une explication réductive, où la réduction peut être totale ou partielle.

⁵ Cette section est largement inspirée de (Jodoin 2013e).

Pour que l'émergence puisse être mobilisée au sein d'une explication il faudrait à tout le moins que le type d'explication retenu soit possible et clairement défini. Sans l'écarter d'emblée, une explication de type inférentiel comme le modèle déductivo-nomologique risque d'avoir maille à partir avec les phénomènes émergents définis à partir de la notion de nouveauté (absente des prémisses). En revanche, une explication de type causal doit conjuguer avec la difficulté de préciser des relata causaux puisqu'il s'agit à prime abord de niveaux d'organisation et non d'entités concrètes, mais avec la difficulté du caractère simultané (synchronique) de ces relata de même qu'avec la réalisabilité multiple.

Par contre, on a vu qu'une explication satisfaisante ne devait pas s'en tenir à l'orthodoxie explicative (Chapitre 2). Face à cette situation théorique présentée plus tôt, celle du projet de la mécanique statistique, et compte tenu des résultats précédents (*ibid.*), trois possibilités se dégagent pour que l'entropie soit un explanandum de l'émergence, pour que l'émergence puisse expliquer l'entropie : (i) une explication au sens faible, où le concept d'émergence vient définir une classe d'entités à laquelle appartiendrait les entités exemplifiant le concept d'entropie ; (ii) une explication en un sens intermédiaire, où l'émergence est entendue dans son acception méthodologique, mettant en lumière cette situation théorique ; (iii) une explication au sens fort, où est défini un élément de la structure effective des processus naturels, ou un processus réel, produisant des phénomènes émergents comme l'entropie.

6.2.1 Quelle entropie ?

Le point de départ qui semble naturel afin d'identifier l'entropie comme un cas d'émergence est sans aucun doute du côté de la version thermodynamique. D'une part, l'émergence, on l'a vu (Chapitre 2), est généralement établie comme un concept relationnel et méréologique où une « base », un niveau d'organisation, généralement l'objet d'une discipline scientifique quasi-autonome, est reliée à un autre niveau d'organisation, objet d'une autre discipline scientifique, et présentant des propriétés et des concepts nouveaux. (Dans des simulations informatiques, on parle plutôt de niveaux de description.) C'est le cas de la thermodynamique et de la mécanique statistique. En effet, et d'autre part, le projet initial de la mécanique statistique était justement de rendre compte de ces nouveaux concepts introduits par la thermodynamique, comme la température (le concept d'énergie est aussi absent de la mécanique newtonienne), nouveaux par rapport à la discipline physique par excellence, la mécanique, mais aussi historiquement, comme l'entropie (Chapitre 3). En ce sens, la mécanique statistique a toujours traité d'émergence – sans le savoir.

Il est donc naturel d'opter pour l'entropie thermodynamique comme relata émergent. Il a été vu que ce concept renvoie à des phénomènes macroscopiques, observables, faisant l'objet de la thermodynamique, une discipline mature et « robuste » qui fait abstraction de la nature et de la structure de la composition des corps macroscopiques. Caractériser ce relata au sein de la taxonomie émergentiste n'est toutefois pas si simple. S'agit-il de l'entropie comme substance, comme propriété ou chose, ou encore comme prédicat, concept ou loi ? En fait, au lieu de s'interroger sur ce type de caractérisation pour le moins fastidieuse, il convient de se pencher d'abord sur l'origine de l'affirmation qualifiant l'entropie, mais aussi le second principe et le principe entropique, d'« émergents ». Car nonobstant la différence entre le contexte scientifique et philosophique d'aujourd'hui et celui du XIX^e siècle où est apparu le concept d'entropie, il est révélateur de constater que ce dernier n'était pas associé à quoi que ce soit d'« émergent ». Ce sont plutôt les développements de la mécanique statistique qui ont suscité ce genre de réflexion, et plus particulièrement, les lacunes, pour ne pas dire les échecs, de son entreprise réductionniste. Ce qui fait défaut ici est de rendre compte de l'irréversibilité thermodynamique (Chapitre 4). Ce n'est pas tant que la mécanique statistique soit incapable de fournir, de manière plus ou moins convaincante, une définition de l'entropie, mais plutôt qu'elle n'arrive pas à reproduire de manière probante l'augmentation d'entropie dans le temps (Chapitre 5). En d'autres mots, elle peut bien offrir un « portrait » statistique de l'entropie, quelque chose de fixe, mais pas un « film », soit le pourquoi et le comment de son évolution temporelle.

Ainsi, ce qui est vraiment émergent, à tout le moins au sens donné par l'émergence explicative, d'une absence de microexplication (stratégie R ; Chapitre 2), est l'irréversibilité, autrement dit l'explanandum que l'on cherche à expliquer par l'émergence, si possible (section 6.2.3). En ce sens, affirmer que l'entropie est une propriété émergente consiste à affirmer que l'une de ses propriétés essentielles, l'irréversibilité, est émergente. Mais toute discussion sur le concept d'entropie, et en particulier en ce qui concerne son caractère émergent, est plus épineuse qu'aux premiers abords. Car si l'émergence est bien un concept relationnel, il y a plusieurs relata crédibles entre ceux de la thermodynamique (macroscopique) et ceux de la mécanique (microscopique). Si l'irréversibilité thermodynamique semble être identifiée comme ce qui émerge (avec les précisions qui s'imposent), la question de savoir à partir de quoi elle émerge est plus difficile, car l'entreprise réductionniste de la mécanique statistique, on l'a vu, a su fournir d'autres définitions de l'entropie. C'est-à-dire qu'elle présente des « images » (Chapitre 2) de l'entropie thermodynamique. En effet, la définition de l'entropie thermodynamique (S_T ; Chapitre 4) est différente de celle de l'entropie statistique de Boltzmann ($S_{B,f}$ ou $S_{B,c}$ ou $S_{B,m}$; Chapitre 5) même si elles peuvent être jugées équivalentes sous certaines conditions. C'est précisément là que se posent les difficultés inhérentes aux relations inter-théoriques, car les critères adoptés pour juger de cette équivalence, autrement dit pour obtenir une

réduction réussie où l'un des concepts ne serait « rien d'autre que » ce qui est signifié par un autre concept, ne sont ni évidents ni consensuels. Puisque ces concepts sont différents, ce qui est trivial s'il s'agit de relations inter-théoriques, il faut justement dire pourquoi il y a équivalence ou pourquoi l'un des concepts se réduit à un autre. Et ce faisant, il faut souvent, en plus, justifier l'emploi de lois-ponts ou une interprétation particulière. Ce n'est qu'à cette condition que les relata et leurs relations peuvent être correctement définis.

En effet, si l'identification d'un cas d'émergence revient d'abord à identifier une « discontinuité conceptuelle » (Chapitre 2), alors plusieurs choix sont disponibles parmi les « images » qu'offre la mécanique statistique du concept d'entropie ou encore dans les interprétations que l'on peut en faire. Prenons l'exemple de l'entropie combinatoire de Boltzmann ($S_{B,c}$), qui est défini selon le logarithme du nombre de complexions ou d'arrangements compatibles avec une distribution ou macro-état donné (Chapitre 5). Dans le cadre théorique de la mécanique selon l'approche de Boltzmann, ce concept aide à définir certaines propriétés (prédicats) des systèmes auquel il est censé s'appliquer. Ces propriétés peuvent être analysées en opposition à celles des constituants, considérés comme élémentaires, de ces systèmes, c'est-à-dire celles des molécules telles que décrites par la mécanique classique (il est aussi possible de le faire avec la mécanique quantique, mais c'est une autre histoire). Sans les énumérer toutes ni apporter des ajustements relativistes, on identifie généralement l'immutabilité et l'impénétrabilité des particules (molécules), les attributs de masse (constante), de vitesse et de quantité de mouvement, et l'invariance par reversement du temps (IRT ; Chapitre 4) des équations du mouvement. Celles-ci sont peut-être discutables, mais le point important est qu'une fois établies, elles servent de relata à l'application des concepts de réduction et d'émergence⁶. Tel que mentionné, elles sont contrastées avec celles d'un contexte empirique particulier, comme celui d'une large collection de molécules, ou d'un autre contexte théorique, comme celui de la mécanique statistique. Par exemple, le nombre d'arrangements dans le concept $S_{B,c}$ est souvent considéré comme définissant la probabilité d'un macro-état et $S_{B,c}$ est souvent interprété comme une mesure du désordre (Chapitre 5). Comme les concepts de probabilités et de désordre sont pour ainsi dire absents de la mécanique classique, d'où la discontinuité conceptuelle, et qu'ils ne peuvent s'y réduire ni être expliqués par cette théorie, alors il est possible de les considérer comme relata émergents et, par voie de conséquence, $S_{B,c}$. Puisque l'entropie statistique est polysémique et qu'elle reçoit plusieurs interprétations, il existe plusieurs candidats comme relata d'émergence par rapport à la mécanique classique, comme base, mais aussi par rapport à la thermodynamique comme émergent.

⁶ Cependant, dans une approche sémantique comme celle de Batterman (2000 et 2002), ces exigences sont plus souples car il est plus question de relations entre théories ou modèles, et non entre entités ou propriétés.

Identifier un concept d'entropie comme explanandum de l'émergence, c'est donc choisir un ensemble de propriétés comme candidats potentiels de relata d'émergence ainsi qu'un ensemble d'hypothèses auxiliaires, mais c'est aussi devoir choisir entre plusieurs interprétations. Il est évident que ce choix vient avec un ensemble de justifications intrinsèques à la définition même de ce concept (par ex. le théorème H), mais aussi d'arguments visant à retrouver les propriétés de l'entropie thermodynamique (par ex. l'irréversibilité). À titre d'exemple, si le choix s'arrête sur l'entropie combinatoire de Boltzmann, comme ce concept fait appel à la méthodologie du grenage grossier pour retrouver la propriété de l'irréversibilité (Chapitre 5), il faut justifier cette méthodologie et faire face à ses critiques (par ex. celle de la subjectivité).

Par ailleurs, le choix d'un concept particulier d'entropie conditionne le type d'émergence susceptible d'en rendre compte. Ainsi, selon Bunge (2003 : 135), une relation micro-macro telle que décrite par l'entropie statistique, et cela vaudrait pour $S_{B,c}$, est une relation de type épistémologique et non ontologique, de sorte que l'émergence ontologique serait exclue dans ce cas précis. Dans le cas de l'entropie de Gibbs ($S_{G,t}$) et de sa méthodologie des ensembles, la question se pose à savoir ce que représente le concept d'entropie ainsi obtenu et conséquemment ce que peut représenter le niveau microscopique et ses diverses relations avec d'autres niveaux. En effet, $S_{G,f}$ est, d'une part, une fonction de distribution sur des ensembles fictifs et non une fonction sur l'espace des phases, et, d'autre part, elle est indépendante de l'évolution du système individuel. C'est pourquoi on peut bien nier, comme Goldstein (2001) qu'elle définisse ou représente l'entropie thermodynamique, ou soutenir comme Lebowitz (1999) qu'elle ne permet pas de comprendre l'évolution temporelle des systèmes macroscopiques. Aussi, le fardeau de la preuve pour en faire un niveau de réalité est sans doute insurmontable. Cette conception élimine donc la relation entre micro- et macro-états et la définition même du niveau microscopique en ces termes semble écarter d'entrée de jeu la possibilité de définir une relation concrète entre entités réelles. En outre, si on ne peut parler de relation ou d'interaction concrète, on peut encore plus difficilement parler de relation de type causale. Comme le fait remarquer Callender (1999), $S_{G,f}$ ne peut être considérée comme la cause de l'entropie thermodynamique parce que ces deux quantités ne covariant pas. Par conséquent, si on ne peut déterminer un niveau inférieur comme niveau de réalité constituant la base d'émergence et une relation suffisamment forte avec ce qui émerge, alors on ne peut parler d'émergence ontologique. Par conséquent, l'entropie de Gibbs ($S_{G,t}$) ne peut exemplifier l'émergence ontologique.

Mais encore une fois, il faut bien préciser les relata d'émergence. Car cette conclusion peut être remise en cause si la base d'émergence n'est pas l'entropie de Gibbs *per se*, mais plutôt les entités de la mécanique (molécules, concepts lois, etc.). De sorte que la tentative d'explication de l'irréversibilité thermodynamique par la méthodologie des ensembles de Gibbs (que l'on jugerait ainsi valide), couplée

à son échec, permettrait d'inférer l'échec d'un certain réductionnisme et par voie de conséquence d'inférer à l'émergence. Cette inférence n'est valide que si la méthodologie des ensembles est valide, qu'elle échoue vraiment, et enfin qu'il n'y pas d'autres réductions disponibles ou que l'émergence soit compatible avec le réductionnisme (celui en cause). Il est question d'un réductionnisme particulier car le réductionnisme générique peut prendre plusieurs formes, en sorte qu'une certaine forme d'émergence est compatible avec une certaine forme de réductionnisme (Wimsatt 1981 et 2006 ; Chapitre 2). Pour une forme de réductionnisme donnée, comme la stratégie R, il faut préciser si les hypothèses auxiliaires servant à l'explication réductive de l'entropie thermodynamique (par ex. la méthodologie du grenage grossier) compte ou non dans l'arsenal autorisé, celui des lois (microscopiques) d'évolution, d'interaction et de composition.

Bien que les différentes versions de l'entropie statistique (Boltzmann et Gibbs) ne rendent pas compte de l'irréversibilité, ou à tout le moins leurs tentatives respectives font face à plusieurs objections, il s'agit de concepts différents et donc de relata émergents potentiellement différents. Ainsi, et surtout, l'entropie de Boltzmann est une fonction du micro-état d'un système individuel tandis que l'entropie de Gibbs est une fonction d'une distribution de probabilité définie sur un ensemble fictif infini de sous-systèmes. Dans ce dernier cas, il n'y a rien qui puisse légitimement représenté une base d'émergence, ni des relata d'une relation de survenance. Dans les deux cas, toutefois, est exemplifiée la réalisabilité multiple au sens où il y a sous-détermination des sous-systèmes compatibles (qu'ils soient « constituants » n'est pas évident avec $S_{G,i}$) avec des paramètres macroscopiques donnés. Si la conception émergentiste de type sémantique, qui présente l'émergence comme une relation entre théories sans définir de relation méréologique, est retenue afin d'éviter la mise de côté du concept gibbsien, alors il faut tout de même préciser ce qui fait partie ou non de ces théories, donc statuer sur ces hypothèses auxiliaires. Il semble que cette tâche se complique car ces hypothèses sont intermédiaires entre la mécanique (et même : le couple mécanique-probabilités) et la thermodynamique.

6.2.2 Quelle émergence ?

Tel que mentionné, il ne fait aucun sens de dire, sans plus de précision, que l'entropie est un cas particulier d'émergence et qu'en même temps elle est ce qui explique l'émergence. L'explication est une notion asymétrique et si, par définition, l'explanans explique l'explanandum, l'inverse n'est pas vrai. Ce qui importe ici est donc l'articulation entre explication et émergence, appliquée au concept d'entropie. Dans un premier temps, l'attention est donc portée ici sur l'entropie comme explanandum

de l'émergence. Dans un second temps, la discussion portera sur l'entropie comme explanans de l'émergence (section 6.3). Alors, si l'entropie est un cas particulier de l'émergence encore faut-il savoir de quel émergence parle-t-on. L'important ici est donc d'identifier un concept d'émergence pouvant légitimement être mobilisé au sein d'un explanans en vue d'expliquer l'entropie. Je présente trois possibilités : l'émergence ontologique, l'émergence interventionniste et l'émergence méthodologique. Bien que l'articulation entre émergence et explication soit primordiale, il a déjà été noté que la catégorie épistémologique de l'émergence explicative est inopérante ici. En effet, elle se définit comme l'impossibilité d'une explication réductive au sens donné par l'impossibilité de la stratégie R (Chapitre 2). Reste à savoir si le phénomène émergent peut tout de même être expliqué, par autre chose qu'une explication réductive, ou par un concept d'émergence défini différemment.

Il importe donc d'identifier le concept d'émergence susceptible de jouer le rôle d'explanans et naturellement on se tourne vers la taxonomie adoptée plus tôt (Chapitre 2). D'abord, l'aspect ontologique du problème. La question revient, pour ainsi dire, à consulter une page de notre catalogue ontologique, à savoir s'il est possible ou non d'éliminer une entité ontologique ou encore de l'identifier à une autre. Tel que mentionné, cet examen est généralement de nature épistémologique, parce que ce qui systématise ou unifie les raisons que l'on peut avoir sur la réalité, du moins telle qu'elle nous apparaît, ce sont les théories scientifiques. Par exemple, pour reprendre la même métaphore, notre catalogue ontologique peut bien comporter les atomes et les molécules ; non seulement les raisons pour lesquelles ces entités s'y trouvent renvoient à des expériences, des modèles et des explications, en bref à l'arsenal épistémique des sciences, mais en plus la réponse à la question de savoir si elles se ramènent d'une quelconque façon à d'autres entités repose sur une investigation profondément théorique : l'affirmation à l'effet que la tasse que je tiens en ce moment soit constituée de molécules ou qu'elle ne soit *que* des molécules trouve sa justification en physique et non en métaphysique. Du moins, c'est par la méthode de la physique et des sciences en général que l'on risque de trouver une justification satisfaisante. D'ailleurs, c'est précisément de cette façon qu'est établie la satisfaction du critère de nouveauté, de « propriété qualitative nouvelle », dans certaine conception d'émergence ontologique du moins. La supériorité des sciences sur la métaphysique n'est pas nécessairement établie *de jure*, mais nous procédons ainsi *de facto*. Néanmoins, la conclusion que l'on peut tirer de ce constat n'est *pas* sans appel pour l'émergence ontologique.

Car, d'une part, malgré qu'une justification soit nécessaire, il est possible d'inférer la portée ontologique d'une certaine entité à partir de son statut épistémique (par ex. que les molécules existent parce qu'elles sont détectables et attestées par plusieurs théories). D'autre part, même si l'on place notre confiance en certaines théories scientifiques pour définir des niveaux d'organisation comme relatifs d'émergence, cela n'empêche pas de construire une théorie émergentiste énonçant l'effectivité d'une

opération physique comme la « fusion des propriétés » (Humphreys 1997 ; Chapitre 2) : une « fusion » consiste en une sorte d'interaction causale singulière où des instances de propriétés à un niveau inférieur se « combinent » pour donner une instance de propriété d'un tout à un niveau supérieur, et où ces instances au niveau inférieur n'existent plus par suite de cette fusion. Une telle opération suffisamment générale peut faire l'objet d'une nomicisation. L'idée générale est de dire « il y a une telle chose dans le monde qu'une opération de fusion entre certaines propriétés à certains niveaux » et donc, à la question « pourquoi telle propriété » la réponse est « parce que telle opération de fusion, donc telle émergence ». Or, dans le cas de l'entropie, il est difficile de dire si les instances de propriétés au niveau inférieur, microscopique, celui des molécules, n'existent plus. On suppose généralement que le micro-état existe, bien que la mesure ne le détermine pas entièrement, mais détermine plutôt une gamme de possibles. D'ailleurs, les différentes expériences de pensée, impliquant des êtres supranaturels comme de démon maxwellien, suppose l'accessibilité empirique du micro-état (Chapitre 5). Par contre, peut-être qu'une intervention directement au niveau moléculaire sur un système macroscopique, une intervention micro-macro, révélerait des surprises (je ne connais pas de telles expériences toutefois).

Par ailleurs, l'entropie peut être considérée comme une propriété émergente selon l'interprétation interventionniste (Chapitre 2). Cette conception se dédouane d'une définition des niveaux de réalité, de relation réelle entre entités concrètes. Il faut toutefois qu'il y ait des propriétés expérimentalement accessibles par une quelconque intervention. Mais observer des molécules avec un instrument tel une chambre à bulles est loin d'exposer les propriétés des entités des niveaux microscopique et macroscopique entrant dans la définition de l'entropie statistique. Il est tout à fait vain ici d'employer un langage elliptique et de prétendre qu'une intervention pourrait s'appliquer directement à une distribution de probabilité ou à un arrangement tels que définis dans le cas de l'entropie statistique. Pour qu'il y ait émergence selon l'interprétation interventionniste, il faut qu'il y ait une intervention au niveau microscopique dont les effets seraient détectés par un mode d'accès approprié et que de manière concomitante des effets au niveau macroscopique puissent aussi être détectés par un mode d'accès ; tel que mentionné, une intervention sur X à un certain niveau d'organisation cause \mathcal{I} , qui est détectable par une autre intervention à un niveau d'organisation inférieur. C'est une condition qui ne peut être remplie si les propriétés sont inobservables par définition, et si elles relèvent d'opérations statistiques (ce n'est pas les valeurs individuelles qui sont en cause, mais leur moyenne), de surcroît pour un ensemble fictif définissant $S_{G,F}$. En effet, dans ce cadre conceptuel, une intervention *définit* les propriétés des niveaux et il n'y a rien à définir sans intervention. C'est aussi une condition rédhibitoire si une intervention au niveau microscopique ne produit aucun changement détectable au niveau macroscopique, et vice versa.

Il devient donc évident que, malgré l'identification de l'irréversibilité comme ce qui émerge, la caractérisation des relata de la base d'émergence, que ce soit des propriétés, des substances ou même des théories (en se soustrayant d'une conception méréologique) est problématique. Il en va de même d'un choix d'un concept d'émergence approprié. Néanmoins, on peut affirmer qu'il y a plusieurs dérivations de concepts exemplifiant la réalisabilité multiple au sens où il y a sous-détermination des sous-systèmes compatibles avec des paramètres macroscopiques donnés, et qui sont eux-mêmes utiles à la prédiction de certaines propriétés thermodynamiques. On peut aussi affirmer que ces concepts ne permettent pas d'explication réductive, au sens dérivational du terme, de l'irréversibilité. Ainsi, malgré des théories scientifiques et des modes d'accès fiables, on se trouve dans l'impasse (en pratique ou en principe) d'en offrir une explication respectant l'orthodoxie. Mais de là à dire que la mécanique statistique et ses concepts d'entropie sont totalement dépourvus de vertus explicatives n'est pas justifié. L'approche de Boltzmann montre que la majorité du volume en phase, aussi considérée comme des conditions initiales, évolue vers des états d'équilibre, d'où le plus grand volume en phase, et toute chose étant égale par ailleurs, il y a plus de chance qu'un système quelconque soit dans un état qui va évoluer vers un état d'équilibre, à condition qu'il demeure isolé (Chapitre 5). L'émergence peut donc aussi être considérée comme un principe *méthodologique* forçant la recherche d'explication réductionniste autant que possible, sans donc multiplier les entités théoriques et sans pour autant coucher les phénomènes sur un lit de mystère. Affirmer qu'un phénomène est émergent c'est affirmer qu'il existe un lien modal (par ex. la survenance) entre des entités à des niveaux d'organisation différents sans qu'une explication réductive ou une réduction totale ne soit possible. C'est ainsi réaffirmer la confiance en des théories opérant à des niveaux d'organisation différents mais entretenant tout de même des liens entre elles, ce que montre bien le concept d'entropie statistique.

6.2.3 Quel explanandum ?

Si un concept d'émergence doit être appliqué à celui d'entropie afin d'évaluer sa pertinence, encore faut-il savoir comment l'appliquer. Car, je le rappelle, ce qui peut être considéré comme émergent peut être l'entropie elle-même, et le cas échéant selon telle ou telle interprétation particulière, mais il est plus pertinent ici d'identifier l'une de ses propriétés, au premier chef la non-décroissance, exemplifiant une irréversibilité (Chapitre 4).

Pourquoi l'entropie serait un explanandum de l'émergence ? Initialement, parce que la mécanique classique s'appliquant aux composants des systèmes thermodynamiques, les molécules, n'est pas en mesure de reproduire l'ensemble des propriétés de l'entropie thermodynamique, au premier chef

l'irréversibilité. Ainsi, l'échec d'une explication réductive, au sens dérivationnel du terme, de l'irréversibilité de l'entropie thermodynamique par la mécanique serait une raison suffisante (croit-on) pour établir cette propriété comme émergente. Mais comme l'irréversibilité semble être une propriété essentielle de l'entropie, on considère (par ex. Sethna 2011) que l'entropie elle-même est émergente ; pour d'autres (par ex. Ao 2007), ce serait la thermodynamique qui serait émergente, avec ses concepts de température et d'entropie.

La discussion précédente a établi que la problématisation de l'émergence de l'entropie implique au moins deux questionnements importants. Il faut en effet définir les relata émergents selon une position épistémologique ou ontologique particulière (substances, concepts, etc.), mais aussi, ce qui est censé s'appliquer, selon une théorie donnée, à cette position épistémologique ou ontologique particulière, comme les propriétés mécaniques des atomes ou une certaine version de l'entropie statistique (en tenant compte ou non des hypothèses auxiliaires). Ainsi, jusqu'à quel point le récit de l'entité réductive (par ex. une théorie) doit-il être « homogène » et ne pas recourir à des entités « extérieures » (par ex. à d'autres théories ou à l'environnement du système en question) ? Par exemple, dans le cas de la prétendue réduction de l'entropie thermodynamique à l'entropie statistique (entropie combinatoire de Boltzmann), est-il vraiment justifié de faire appel au calcul des probabilités pour décrire le comportement des molécules dont les propriétés intrinsèques semblent être de nature mécanique et non pas probabiliste ? Autrement dit, est-ce que les probabilités constituent un élément justifié des lois-ponts ? Car s'il est effectivement justifié de procéder à une réduction par le truchement des probabilités, alors la réduction peut être considérée comme réussie, éliminant ainsi certains types d'émergence, et dans le cas contraire, ce serait plutôt une bonne raison d'affirmer l'émergence de ce qui ne peut être réduit.

Il a été vu aussi que, dans son formalisme, la mécanique statistique stipule qu'il y a deux niveaux, l'un macroscopique et observable, déterminé par les propriétés thermodynamiques, et l'autre microscopique et inobservable, déterminé par les propriétés dynamiques (Chapitre 5). Cette hiérarchie de niveaux entraîne l'exemplification de la réalisabilité multiple, car plusieurs micro-états sont associés à un seul macro-état, laquelle et on l'interprète souvent comme exprimant une probabilité ou le désordre. Ce qui est à l'étude ici est donc le concept d'entropie comme un cas d'émergence synchronique. Le récit standard expliquerait ainsi l'augmentation d'entropie par un certain principe voulant qu'un système ait tendance à passer d'un état peu probable à un état plus probable. Cependant, on a plutôt établi qu'un système thermodynamique isolé a une forte probabilité synchronique d'être dans un micro-état évoluant de manière déterministe vers un micro-état associé à une entropie statistique élevée. Il a donc un niveau synchroniquement sous-déterminé dont la caractérisation (partielle) repose sur des paramètres, déterminés par la mesure (empirique), lesquels

circonscrivent un domaine de variations des propriétés des éléments de ce niveau sous-déterminé à la manière de conditions limites.

Alors, nonobstant les difficultés mentionnées relatives aux relata émergents, la question se pose de savoir si la propriété d'entropie est une *propriété structurale*, à l'instar de la forme que possède une pyramide de briques, par exemple. Et il faut statuer sur ce niveau de propriétés structurales : micro-, méso-, macroscopique ? En effet, les définitions statistiques de l'entropie font référence à des propriétés structurales en ce sens qu'elles se révèlent dans l'organisation de grands ensembles (statistiques) de molécules et non chez les entités qui composent ces ensembles. On parle d'ailleurs d'arrangements dans l'espace des phases : de distribution d'états dynamiques ($S_{B,f}$ ou $S_{G,f}$), d'arrangement combinatoire ($S_{B,c}$) ou de volume en phase ($S_{B,m}$). Un des problèmes avec les propriétés structurales est qu'elles présupposent un niveau plus fondamental à partir duquel elles surviennent sans être pour autant émergentes, comme quoi elles ne peuvent constituer un niveau d'organisation ou de réalité. Cela peut bien constituer un problème, en général, dans l'explication de systèmes complexes et des niveaux d'organisation (Chapitre 2), mais une telle conception semble tout de même idoine au contexte de la mécanique statistique. Or, s'il y a bien un niveau qui puisse être considéré comme fondamental, il semble bien que ce soit celui de la mécanique appliquée aux atomes et molécules.

Par contre, dans le cadre de la mécanique statistique, les molécules comme leur arrangements structurales ne sont pas observables. Ce n'est pas qu'ils soient inobservables en principe, mais dans ce contexte la valeur des variables mécaniques ne peut être déterminée, seul le domaine de variations. Cette motivation d'expliquer l'observable par l'inobservable repose tout de même sur ce qui semble indubitable et fondamental, soit la mécanique classique, même si le calcul statistique intervient. Il est alors possible de mesurer l'ampleur des tensions épistémologiques potentielles, car ce qui peut paraître indubitable pour l'un ne l'est pas nécessairement pour l'autre : un empiriste favorisera ainsi, pour une justification ou une explication, la thermodynamique, tandis qu'un réaliste pourrait opter pour la mécanique. Cependant, selon le critère de robustesse, où une entité doit être accessible par une multitude de modes d'accès indépendants, il semble que la thermodynamique soit avantagée par son accessibilité empirique. Au premier chef, les états dynamiques exemplifient la réalisabilité multiple, qui est un cas de sous-détermination et non de détermination multiple ou de surdétermination ; de plus, ils ne sont pas empiriquement accessibles et encore moins observables au sens strict. Du point de vue empiriste, ce sont les phénomènes observables, ici macroscopiques et thermodynamiques, qui constituent une base d'inférence indubitable ou à tout le moins un objet de description épistémologiquement fondée. Du point de vue réaliste, ceux-ci constituent un ensemble robuste. Ainsi, d'une part, la stratégie explicative de la mécanique statistique articulant le concept d'entropie tenterait

d'expliquer ce qui est indubitable par ce qui ne l'est pas. Il y aurait un renversement de la stratégie, exposée par Wimsatt (2012), voulant que l'on explique ce qui n'est pas robuste par ce qui est robuste.

6.2.4 Discussion & conclusion

Malgré que le statut épistémologique du concept d'entropie dans sa formulation statistique ne soit pas clair et qu'il ne soit pas évident que l'une de ses différentes versions parvienne à établir un statut épistémique fort, comme une preuve, il demeure un concept fondamental et la mécanique statistique offre une explication (plus faible) du concept d'entropie thermodynamique. En effet, d'une part, il y a plusieurs sens selon lesquels l'entropie peut être considérée comme un cas particulier d'émergence ou encore que l'émergence puisse expliquer l'entropie. D'autre part, même si l'entropie devait être un cas d'émergence, la stratégie explicative de la mécanique statistique ne serait pas pour autant nécessairement vaine ou inopérante.

Le modèle d'explication présenté plus tôt (Chapitre 2) suggère ainsi d'emprunter la notion d'espace des phases afin de circonscrire la gamme des possibles pour un explanandum donné. En identifiant d'abord une région de l'espace des phases, sont définies deux classes d'énoncés contrastives circonscrivant ce qui est possible et ce qui ne l'est pas. Ensuite, il faut appuyer un lien modal entre l'explanans et l'explanandum. Les critères pertinents de l'orthodoxie d'une « bonne » explication, comme les lois, les causes et les régularités statistiques permettent de soutenir une expectabilité, de répondre à ces schèmes systématiques de dépendance contrefactuelle, mais exemplifient aussi les concepts de modalité (au sens faible) et de naturalité. De façon générale, une explication doit donc montrer que ce qui s'est passé (explanandum) pouvait (gamme de possibles) se passer, plutôt que tel autre explanandum (impossible) ; ou que cela devait, selon une modalité faible, se passer ; ou que dans des circonstances différentes cela se serait aussi, naturellement, produit. De la discussion précédente (section 6.2.1 à 6.2.3), trois possibilités s'ouvrent pour l'entropie comme explanandum de l'émergence, soit dans le schème d'analyse conceptuelle où l'émergence (en un sens *A*) explique (en un sens *B*) l'entropie (en un sens *C*).

D'abord, la définition de deux classes d'énoncés contrastives circonscrivant ce qui est possible et ce qui ne l'est pas offre une explication au sens faible. La question du « degré d'explication » (fort ou faible) ne sera pas traitée en profondeur ici, mais elle peut s'illustrer (je crois) comme l'écart séparant une explication dérivationnelle (inférentielle), comme dans un cas déterministe (absolu), d'un autre type d'explication où le lien entre explanans et explanandum n'est pas aussi fort (nonobstant quelques difficultés pouvant être rattachées aux explications unificationniste et pragmatique). De plus, cette

notion de « degré d'explication » reprend bien (je crois) cette intuition voulant que certaines explications (par ex. du sens commun) soient valables même s'il y en a d'autres jugées meilleures (par ex. scientifiques). Maintenant la définition d'un concept adéquat d'émergence, respectant le critère de non-trivialité, permet de définir (trivialement) une classe de référents pouvant s'y appliquer et une classe de référents ne le pouvant pas. Pour un ensemble de référents donnés (par ex. la classe des systèmes thermodynamiques), il est possible d'établir si tel concept d'émergence s'y applique ou pas. Dire ainsi que l'entropie macroscopique de Boltzmann, par exemple, est émergente au sens où elle présente une « propriété qualitative nouvelle » (structurale ou non), délimite déjà une classe d'entités pouvant en faire partie. Par exemple, dire que telle entité appartient à la classe « cheval » permet ainsi de situer conceptuellement cette entité en lui attribuant un certains nombre de propriétés qui, ensemble, définissent cette classe plutôt qu'une autre.

Cependant, il est évident que ce genre d'explication est faible, et dans plusieurs cas insatisfaisant, puisqu'elle s'apparente à nommer ou stipuler, *per vim definitionis*, à l'instar des « vertus dormitives » de l'opium. Elle s'en distingue toutefois, car elle opère sur un fond de connaissances acquises qui permet, justement, d'établir certaines propriétés attribuées aux référents potentiels, de circonscrire le domaine de possibles pour ces référents, son espace des phases en quelque sorte. Pour reprendre le même exemple, le prédicat « n'est pas un reptile » fait partie implicitement du concept de « cheval » et un fond de connaissances permet d'établir qu'un prédicat comme « résiste à des températures supérieures à mille degrés Celsius » ne s'y applique guère. Ainsi, dire que la propriété ou le concept d'entropie est émergent permet aussi de circonscrire, parfois vaguement certes, un ensemble de prédicats pouvant s'y appliquer.

Ensuite, une explication en un sens intermédiaire est aussi possible, où l'émergence est entendue dans son acception méthodologique (Chapitre 2). L'émergence méthodologique encourage la recherche d'explication réductionniste autant que possible, sans toutefois croire qu'elle soit toujours possible ou souhaitable, et sans freiner la systématisation des connaissances en multipliant les entités pertinentes à l'explication. Affirmer qu'un phénomène est émergent c'est affirmer qu'il existe un lien modal entre des entités à des niveaux d'organisation différents. La survenance, où il n'y a pas de différences à un niveau inférieur sans différence à un niveau supérieur, est un exemple. Mais ce lien est plutôt celui entre constituants et constitué, comme dans le cas des atomes et des systèmes thermodynamiques. C'est d'ailleurs en ce sens que, dans l'exemple précédent, la survenance des niveaux fait vraiment sens. C'est aussi en ce sens que l'émergence est un concept relationnel excluant une émergence *ex nihilo*. Or, ce lien modal ne garantit pas qu'une explication réductive au sens dérivationnel du terme soit possible. C'est-à-dire ici que la mécanique statistique ne peut offrir de preuve de l'augmentation d'une fonction de l'espace des phases censée représenter l'irréversibilité thermodynamique. En ce sens, identifier un cas

d'émergence revient à caractériser une situation théorique particulière où différents modes d'accès permettent de définir des niveaux d'organisation sans qu'il n'y ait de réduction totale (identité ou élimination).

L'émergence méthodologique propose aussi de suspendre son jugement lorsque plusieurs modes d'accès définissent des niveaux d'organisation différents sans que leur valeur épistémologique ne soit clairement établie. Cet exemple de la mécanique statistique et de la thermodynamique est à cet égard probant. Tel que mentionné, il y a un renversement, en un sens assez clair, de la stratégie exposée par Wimsatt (2012) voulant que l'on explique ce qui n'est pas robuste par ce qui est robuste. D'un côté on retrouve une théorie formelle s'appuyant sur les lois fondamentales de la mécanique appliquées à ce qui est considéré comme les constituants de la matière mais qui demeurent inobservables, du moins leur valeur instantanée (et l'on pourrait ajouter qu'elle n'est qu'une approximation⁷ de la « vraie » mécanique, quantique cette fois). D'un autre côté on a la thermodynamique, reposant sur des principes généralisant et étant corroborés par un nombre incalculable d'observations, et donc le comportement des systèmes qu'elle décrit est observable, mais sans doute un peu moins mathématisée et moins orthodoxe par rapport à la plupart des théories physiques (comme le parangon qu'est la mécanique, qui ne manifeste pas un besoin aussi grand de faire référence, entre autres, à l'environnement des systèmes étudiés). Ainsi, même si une réduction, qui risque d'être partielle plutôt que totale, était possible, les vertus cognitives ou représentationnelles de la thermodynamique demeureraient néanmoins, et il ne serait pas absurde d'affirmer qu'elle est aussi établie, qu'elle possède des vertus épistémologiques aussi grande que la mécanique. L'émergence méthodologique réaffirme la valeur de ces vertus épistémologiques face à la valeur d'une réduction, qui demeure néanmoins désirable.

Enfin, il est possible d'offrir une explication au sens fort, en ce sens où est défini un élément de la structure effective des processus naturels, ou un processus réel, produisant des phénomènes émergents comme l'entropie. Cela revient à caractériser la relation entre entités à des niveaux différents de manière causale ou du moins selon un lien modal fort. Évidemment, cette caractérisation sera très certainement une catégorie métaphysique. Il s'agit de l'une des raisons pourquoi un cas d'émergence peut aussi être expliqué, c'est-à-dire que les critères définissant l'émergence ne sont pas en premier lieu épistémiques (Chapitre 2).

Le schème formel où l'émergence (en un sens *A*) explique (en un sens *B*) l'entropie (en un sens *C*) se décline donc en trois possibilités. Sans reprendre l'ensemble de la discussion précédente, on peut noter que le sens *A* peut être celui d'une espèce d'émergence générique définissant deux classes d'énoncés

⁷ On rappelle toutefois que l'affirmation à l'effet qu'une théorie constitue une approximation d'une autre demande de la subtilité et ne peut être avancée à la légère (voir Chapitre 2).

contrastives, celui de l'émergence méthodologique ou celui de l'émergence ontologique (causale ou non). Le sens de B dépend de l'émergence choisie (donc du sens de A), tandis que l'entropie (sens C) est l'entropie thermodynamique, nonobstant quelques subtilités abordées plus tôt (sections 6.2.1 à 6.2.3). Ainsi, affirmer l'émergence de l'entropie n'élimine pas la possibilité d'une explication en général ni même d'une explication de type réductif. Et si l'entropie n'est pas une propriété émergente, alors deux possibilités s'ouvrent, un macro-réductionnisme ou un micro-réductionnisme : soit elle constitue une propriété macroscopique robuste et autonome par rapport aux constituants des systèmes auxquels elle s'applique, en conséquence de quoi elle ne peut s'appliquer à et encore moins expliquer des niveaux sous-jacents ; soit elle constitue une propriété microscopique pouvant expliquer des phénomènes à différents niveaux d'organisation, mais sans base empirique solide, sans robustesse, et sans justification de son augmentation. Deux avenues guère prometteuses.

6.3 L'entropie comme explanans

L'émergence explique l'entropie thermodynamique en un sens faible (subsumption sous un concept définissant une classe d'émergents) et intermédiaire (émergence méthodologique), alors qu'une explication au sens fort demande tout un programme métaphysique (émergence ontologique), ou encore une preuve ou une dérivation réductive. Si ce dernier cas de figure se révèle impossible, on considère, selon un point de vue restreint cependant, que l'entropie thermodynamique serait une propriété émergente des propriétés des entités que sont les molécules régies par la mécanique, mais sans en fournir d'explication car elle est synonyme dans ce cas de non-explicabilité (émergence explicative). Ainsi, devant les tentatives au succès mitigé de la mécanique statistique de réduire le concept d'entropie thermodynamique à un concept formulé en termes de variables mécaniques et statistiques, il est souvent soutenu que le concept d'entropie thermodynamique, et plus particulièrement la propriété d'irréversibilité, est un cas d'émergence. En revanche, l'entropie thermodynamique peut faire partie d'un explanans expliquant, au sens déductivo-nomologique ou causal-mécanique, un phénomène d'émergence, au sens de l'apparition d'une propriété qualitativement nouvelle, d'une non-élimination ou d'une non-identité (Chapitre 2). L'entropie expliquerait, *mutatis mutandis*, cette émergence. En ce sens l'entropie serait un explanandum de l'émergence. C'est précisément là que se présente le paradoxe, du moins en apparence, car cela signifie que le concept d'entropie serait à la fois explanandum et explanans de l'émergence.

Je distingue ici deux approches quant au rôle du concept d'entropie au sein d'un explanans de l'émergence. D'une part, l'*approche substantielle* où le concept d'entropie renvoie à la structuration effective des processus naturels qui s'accompagne d'une charge modale, une force de changement pouvant mener à l'apparition et éventuellement à l'émergence de certaines propriétés. En d'autres termes, sans s'y limiter cependant, l'entropie aurait un pouvoir causal. En ce sens, cette approche s'appuie sur le contenu de signification du concept d'entropie appliqué à des systèmes divers (comme des systèmes biologiques) mais considérés comme des systèmes soumis spécifiquement aux lois physiques. D'autre part, l'*approche analogique* s'appuie sur une ou des analogie(s) formelle(s) entre la structure ou des propriétés mathématiques de l'entropie en physique statistique et celle de la formalisation des processus biologiques évolutionnistes. La stratégie commune de cette approche est d'identifier une population de sous-unités d'un large système comme éléments d'un espace des phases servant à définir un concept spécifique d'entropie et d'y attribuer la propriété de non-décroissance de l'entropie.

L'idée générale derrière le rôle du concept d'entropie au sein d'un explanans de l'émergence de certaines propriétés est, selon l'approche substantielle et en termes simples, la suivante : lorsqu'un système suffisamment complexe reçoit une certaine quantité d'énergie, il adopte une configuration particulière, qui « émerge », alors qu'il consomme une quantité de travail et se débarrasse d'une certaine quantité d'entropie. Plusieurs éléments de cette description servent ici de paramètres : la complexité du système (ce qui signifie qu'il a suffisamment de degrés de liberté et présente une certaine adaptabilité, une capacité de changer), la configuration (une structure ordonnée, une organisation particulière, un degré de complexité ?) qui est censée émerger (en quel sens ? la nouveauté qualitative, l'auto-organisation ?), et la concomitance (est-elle modale ? peut-elle servir d'explication ?) de cette émergence avec ce processus de consommation de travail et de dissipation d'entropie. Bref, plusieurs questions sont à l'ordre du jour. Mais cette stratégie permet en principe d'expliquer plusieurs phénomènes divers et poursuit ce bouleversement de l'« *a priori* de la biologie » (Jacob 1970 : 212) séparant les êtres des choses, d'une part en les embrassant au sein de la classe des systèmes thermodynamiques et, d'autre part, en montrant que l'organisation de certains de ces systèmes est régie par des principes physiques.

La stratégie couramment employée dans l'approche analogique a déjà été présentée, mais des exemples sont ici nécessaires. Fisher (1922 et 1930) compare une population de facteurs mendéliens (gènes) à un ensemble de molécules constituant un gaz où la multitude des causes individuelles peut être négligée et l'attention portée sur le comportement statistique de l'ensemble (Chapitre 3). Grâce à cette analogie, Fisher est en mesure, d'une part, de concevoir une population idéalisée lui permettant de déterminer des effets qui ne seraient pas mesurables expérimentalement et, d'autre part, de calculer la distribution de fréquence des facteurs mendéliens (gènes) sur la base d'une distribution stable en l'absence de sélection, drift, etc. (à l'instar de la distribution de vitesses des molécules dans un gaz). Les modèles en génétique des populations de Sella & Hirsch (2005) et de Barton & Coe (2009) seront ici discutés. Un autre exemple vient de Demetrius (2000), pour qui la théorie de l'évolution et la thermodynamique sont deux disciplines où la structure mathématique comporte un principe décrivant un accroissement unidirectionnel d'une propriété opérationnellement mesurable. Selon lui (voir aussi ses 1974 et 1977), la *fitness* darwinienne est entièrement décrite par l'entropie évolutionnaire, une variable démographique qui mesure l'hétérogénéité dans l'âge d'une population d'individus se reproduisant entre eux. La question est donc de savoir s'il est légitime, et sous quelles conditions, de transposer ce qui fait le succès de la physique en biologie ?

Ces deux approches mobilisent donc des concepts d'entropie différents et de façons différentes. Alors que la première adopte une stratégie causale (encadrée par des lois ou principes), la seconde infère déductivement l'explanandum à partir d'une loi ou d'un principe (en l'occurrence le principe

entropique) et de la correspondance formelle d'éléments semblables entre, d'une part, l'explanandum en question, et d'autre part, les explananda propres à cette loi ou principe. Dans les deux cas, cependant, il est question de la légitimité de la généralisation de l'application d'un principe à des systèmes qui, au demeurant, ne sont pas propres à la théorie encadrant ce principe et la description de ces systèmes.

6.3.1 Approche substantielle de l'entropie comme explanans

La thermodynamique est née d'observations relativement simples ayant mené à l'élaboration de principes fondamentaux, comme l'impossibilité du mouvement perpétuel, et à des concepts tout aussi fondamentaux, comme ceux d'énergie et d'entropie (Chapitre 3). Le concept d'entropie devait aussi jouer un rôle important dans l'explication de ces phénomènes observés. Le principe entropique stipule que l'entropie globale (système et environnement) ne peut qu'augmenter, en sorte que la chaleur passe spontanément du chaud vers le froid parce que l'inverse impliquerait une diminution (impossible) de l'entropie globale (Chapitre 4). La chaleur peut passer indirectement du froid vers le chaud, mais pas de manière spontanée, plutôt par le truchement d'une intervention sur le système par l'investissement d'une quantité de travail, qui implique une diminution locale d'entropie et une augmentation globale d'entropie (*ibid.*). Le concept d'entropie est donc mobilisé au sein d'explication de systèmes relativement simples, typiquement thermodynamiques (macroscopiques et proches de l'équilibre), qui sont décrits d'après un petit nombre de variables comme le volume, la température et l'énergie. Il est moins évident toutefois qu'il puisse s'appliquer à des systèmes plus complexes, comme des systèmes biologiques. Ces derniers sont aussi des systèmes physiques et en ce sens le concept d'entropie s'y applique de manière plutôt triviale, car ils ne violent pas le principe entropique. Mais que ce dernier puisse vraiment jouer un rôle positif dans l'explicatif des phénomènes biologiques ne va pas de soi et il faut alors le généraliser aux cas de non-équilibre. L'idée a toutefois gagné en crédibilité et se place désormais au centre de plusieurs stratégies explicatives (par ex. Weber *et al.* 1989 ; Schneider & Kay 1994 ; Margulis & Sagan 2000 ; Sheehan 2007 ; Coffman 2011).

L'application du concept d'entropie aux systèmes complexes pose de nombreux défis. D'abord, du point de vue de la thermodynamique, comme ces systèmes sont généralement et typiquement en interaction avec leur environnement, ils seront rarement dans un état d'équilibre, tout au plus dans un état stationnaire. Il est donc difficile, sinon impossible de leur attribuer une valeur de température bien déterminée, nécessaire à la définition thermodynamique de l'entropie. Ensuite, du point de vue de la mécanique statistique, comme ces systèmes sont formés de sous-systèmes en interaction mutuelle et

selon différents niveaux d'organisation, l'hypothèse de l'indépendance statistique des composants, nécessaire, du moins pour des raisons pratiques compte tenu du nombre gigantesque d'interactions possibles, à la définition statistique de l'entropie, ne tient plus. Mais un défi ne constitue pas un obstacle insurmontable.

6.3.1.1 STRATÉGIE GÉNÉRALISÉE ET SYSTÈMES COMPLEXES

La « stratégie généralisée » du principe entropique est la suivante (Chapitre 4) : (i) toute transformation implique des échanges d'énergie et tout échange d'énergie implique une augmentation totale d'entropie, soit une diminution de la capacité d'un système isolé à effectuer un travail ; (ii) cette capacité est interprétée comme la possibilité d'obtenir un « niveau d'ordre », une « structure ordonnée », une « complexité organisée » ; (iii) enfin, étant donné cette tendance, cette direction des processus isolés vers l'augmentation d'entropie, cette capacité du système à effectuer un travail peut être maintenue par « surcompensation » dans un système ouvert, c'est-à-dire par une interaction avec son environnement de manière à diminuer l'entropie du système et à augmenter celle de l'environnement. Dans un contexte d'échanges, de flux importants de matière et d'énergie, le point crucial est d'associer certaines quantités échangées entre le système et son environnement à l'entropie de sorte que ce système soit considéré loin de l'équilibre, c'est-à-dire loin de sa valeur maximale d'entropie, et donc (selon une interprétation particulière) d'un état de « désordre ». Dans le cas d'une voiture, par exemple, parmi ces quantités se trouve la chaleur dégagée par le moteur, mais aussi la matière comme des gaz d'échappement ou même les traces de caoutchouc laissées par les pneus de la voiture, lesquelles quantités sont alors traduites en quantités d'entropie.

L'idée derrière cette généralisation consiste en ce qu'un système subissant des transformations d'énergie voit son entropie augmenter à moins d'entrer en interaction avec son environnement de manière à la diminuer, et ce faisant, de nouvelles « structures » (à définir), dont certaines « ordonnées » (à définir), sont disponibles à la composition générale de ce système. Une telle généralisation joue souvent sur le plan de la métaphore où, d'une part, la matière et l'énergie pouvant être utilisées ou exploitées par le système afin de construire, d'augmenter le « niveau d'ordre », sont associées à une capacité d'effectuer un travail dit « utile », et où, d'autre part, la matière et l'énergie ne pouvant être exploitées par le système, généralement sous forme de chaleur, bref des « déchets », sont considérées comme de l'entropie. Cette interprétation vient elle-même de l'interprétation de la chaleur comme forme d'énergie « dégradée » (Chapitre 4). Sans une interaction particulière entre le système et son environnement, certains états sont inaccessibles au système et cet état de fait est attribué à une augmentation d'entropie. Ainsi, certains processus des systèmes complexes ne sont possibles *que* parce qu'ils exploitent l'énergie de leur environnement tout en y rejetant de l'entropie, en bref par des flux de

matière, d'énergie et d'entropie gardant le système loin de l'équilibre : « les processus de la nature complexe et active, notre propre vie, ne sont possibles que parce qu'ils sont maintenus loin de l'équilibre par les flux incessants qui les nourrissent » (Prigogine & Stengers 1979 : 265).

La conceptualisation de l'entropie comme élément d'un explanans de l'émergence apparaît ici assez clairement : des flux d'entropie particuliers seraient la *cause* de l'émergence de certaines propriétés, de certaines structures ou configurations. Bien entendu, il y a de nombreux défis à définir ce qu'est une cause (Chapitre 2). Ici aussi, la théorie interventionniste de la causalité paraît suffisante, où le caractère explicatif des inférences causales repose essentiellement sur la dépendance contrefactuelle d'une manipulation d'une variable antérieure (cause) au changement d'une autre variable (effet), le but étant de caractériser de manière générale la stratégie explicative. Les caractéristiques de cette cause sont tout de même nomologiquement encadrées par le principe entropique.

Nicolis & Prigogine (1977) proposent le terme « structure dissipative » pour désigner un système capable d'exploiter un apport d'énergie de basse entropie (haute exergie) afin de se maintenir loin de l'équilibre en se débarrassant d'une certaine quantité d'entropie. Or, ce qui est considéré comme émergent est généralement des propriétés typiques des systèmes complexes comme un haut degré d'organisation, une certaine organisation structurale. Il faut toutefois répondre à une question à laquelle on ne peut échapper, qui est celle de l'explanandum, à savoir si ce qui est expliqué est bien l'*émergence* de certaines propriétés ou seulement leur *apparition*. Il faut aussi faire preuve de prudence avec une telle interprétation entropique de type causaliste de l'explication des phénomènes de complexité. On rappelle que le principe entropique restreint la gamme des possibles des phénomènes énergétiques : parmi les processus respectant le principe de conservation de l'énergie, sont possibles uniquement les processus compatibles avec une augmentation globale d'entropie. En ce sens, on peut se demander si ces flux d'entropie agissent bien de manière causale, du moins s'ils peuvent être décrits tels au sein de l'explanans, ou plutôt si le principe entropique ne vient pas seulement imposer des contraintes aux limites sélectionnant certains états possibles. Une façon de l'interpréter, qui ne présume pas de sa validité, mobilise le concept de causalité formelle.

Empruntant aussi à la dynamique, on peut interpréter l'augmentation d'entropie au sein du principe entropique par rapport au concept d'attracteur, comme un ensemble d'états possibles, donc une région de l'espace des phases, vers lequel tend à évoluer le système soumis à une certaine dynamique. Un état d'entropie élevée, un état d'équilibre, représenterait un tel attracteur. En ce sens le principe entropique affirme l'existence d'un état final, l'état d'équilibre. C'est pourquoi, sans toutefois les cautionner, certains y voient une « cause finale » (Slathe 2004) ou le reflet d'une « téléologie naturelle » (Weber 2007). Mais ces expressions sont trompeuses. Nul besoin toutefois de faire appel à ce

genre d'expressions connotées. Car elles occultent le fait que les processus menant à l'état final sont circonscrits par les lois de la physique et qu'il est supposé que leur sous-structure soit déterministe selon la mécanique classique. Ensuite, l'idée de cause finale exprime le dessein de l'existence d'une chose ou l'intentionnalité dans la réalisation d'un but par des moyens qui dépendent de manière projective de ce même but. Enfin, elle confond la condition nécessaire ou la conséquence d'un énoncé conditionnel avec la notion de cause. À moins de revisiter l'entièreté de la physique moderne, en rejetant par exemple les équations différentielles, la notion de cause finale n'y a pas sa place.

De plus, une diminution locale d'entropie, compatible avec une augmentation globale, n'entraîne pas nécessairement un haut degré d'organisation, à moins d'y adosser une interprétation suffisamment indulgente. L'indulgence peut aller jusqu'au point où un état organisé est *définie* comme un état de faible entropie. Mais évidemment, une explication *per vim definitionis*, autrement dit une stipulation à l'instar des « vertus dormitives », n'explique rien – à moins de situer un phénomène dans un cadre théorique bien établi.

La stratégie généralisée du principe entropique des systèmes complexes, ouverts, exige donc un flux continu de matière et d'énergie ainsi qu'une dissipation d'entropie dans leur environnement. En tant que tel, il n'y a rien de particulier ou d'intéressant, il faut plus. En fait, ce qui suscite l'intérêt ce sont certains traits caractéristiques de ces systèmes complexes présentant une telle structure dissipative alors qu'ils sont loin de l'équilibre. C'est précisément cette structure, qu'il reste à préciser, qui est censée être une propriété émergente. L'idée générale est donc que les phénomènes émergents sont rendus possibles par un comportement particulier de l'entropie. Plus spécifiquement, le niveau ou l'intensité de l'émergence, qui demeure à définir et qui en première approche s'apparente à la complexité et à l'organisation, d'un système donné, covarie avec la valeur de l'entropie de ce même système. Encore plus spécifiquement, le mécanisme à l'œuvre dans les phénomènes émergents est souvent décrit comme l'expulsion ou le rejet d'une quantité d'entropie du système considéré comme émergent, en conséquence de quoi l'entropie de l'environnement doit augmenter. Et une condition supplémentaire est celle voulant que le système soit en consommation continue d'énergie dans un état loin de l'équilibre.

Cette interprétation de l'entropie comme explanans de l'émergence n'est pas surprenante si l'on conçoit l'entropie comme une « mesure du désordre » et l'émergence comme étant définie par l'apparition d'un ordre « nouveau », de sorte que si l'un diminue l'autre augmente, et vice versa. Par exemple, on affirme que lorsque des systèmes complexes hautement ordonnés « émergent », ils se développent et évoluent aux dépens d'un « désordre » croissant aux niveaux supérieurs de la hiérarchie (Schneider & Kay 1994). Ainsi, la diminution locale d'entropie « expliquerait » l'apparition d'une

« nouveauté » en général, d'une structure particulière, d'un « ordre » ou d'une « complexité ». Pour Mainzer (2005 : 154) « l'émergence de l'ordre est rendue possible par les transitions de phase de systèmes ouverts interagissant avec leur environnement »⁸, autrement dit en rejetant leur entropie dans l'environnement. Et pour Schneider & Kay (1994 : 31), « l'émergence de la structure ordonnée résulte en une production d'entropie accrue »⁹. Ou encore « le ratio du taux de la génération d'entropie et de l'entropie propre à une structure par rapport à leur environnement peut être considéré comme un indicateur de la capacité d'un système dissipatif s'auto-organisant à se maintenir loin de l'équilibre en rejetant son entropie »¹⁰ (Ludovisi, Pandolfi & Tattichi 2005). Malgré les précautions que demande ce genre de stratégie explicative, on peut donc concevoir que les vertus explicatives du concept d'entropie sont potentiellement très grandes. Par exemple, en présentant le principe entropique comme l'explans de l'origine de la complexité locale, de l'organisation de tout système suffisamment complexe et ouvert sur son environnement, c'est en définitive le phénomène du vivant qui pourrait être ainsi expliqué.

Un exemple est de mise et celui des cellules de Rayleigh-Bénard est notoire. Dans certaines conditions impliquant une fine couche de liquide (comme l'eau) et des températures uniformes aux surfaces inférieure et supérieure, lorsqu'est chauffée (apport d'énergie) la partie inférieure et que la différence de température avec la partie supérieure atteint un seuil critique (un certain nombre de Rayleigh), le système s'éloigne de l'équilibre et manifeste une certaine instabilité. Une structure ou configuration particulière apparaît alors, des cellules de convection hexagonales commencent à se former – les cellules de Bénard. Cette structure a cette particularité d'augmenter la dissipation de la chaleur (ici dans la partie supérieure) et donc de limiter le gradient de température entre les parties inférieure et supérieure. Lorsque l'apport d'énergie cesse et que le gradient de température diminue, la configuration déterminée par les cellules de convection disparaît. Ici, l'émergence serait celle de la configuration, de la structure dissipative (hexagonale) des cellules de convection, et la condition serait le rejet d'entropie dans l'environnement. Il y a aussi un autre élément important à souligner, soit que le système consomme une certaine quantité de travail pour la formation de cette structure ou configuration.

Trois éléments sont donc identifiés dans la stratégie généralisée du principe entropique : (i) un apport d'énergie de l'environnement vers le système, (ii) la consommation d'un travail menant à la

⁸ « *The emergence of order is made possible by phase transitions of open systems interacting with their environment.* »

⁹ « [...] *the emergence of the ordered structure results in more entropy production.* »

¹⁰ « *The ratio of entropy generation rate to entropy embodied in structures relatively to the surroundings can be considered as an indicator of the ability of a self-organizing dissipative system to maintain itself far from equilibrium by pumping out entropy.* »

formation d'une structure ou d'une configuration particulière, et (iii) la dissipation ou le rejet d'une quantité d'entropie dans l'environnement. L'idée générale derrière le rôle du concept d'entropie au sein d'un explanans de l'émergence de certaines propriétés est alors la suivante : lorsqu'un système suffisamment complexe reçoit une certaine quantité d'énergie, il adopte une configuration particulière, qui « émerge », alors qu'il consomme une quantité de travail et se débarrasse d'une certaine quantité d'entropie. Tel que mentionné plus tôt, plusieurs éléments de cette description servent ici de paramètres : la complexité du système (signifiant qu'il ait suffisamment de degrés de liberté et présente une certaine adaptabilité, une capacité de changer), la configuration (une structure ordonnée, une organisation particulière ?) qui est censée émerger (en quel sens ?), et la coexistence ou de la concomitance (est-elle modale ? peut-elle servir d'explication ?) de cette émergence avec ce processus de consommation de travail et de dissipation d'entropie. Bref, plusieurs questions sont à l'ordre du jour. Ce n'est pas étonnant compte tenu des prétentions explicatives de cette stratégie. Déclinons cette dernière.

STRATÉGIE GÉNÉRALISÉE DU PRINCIPE ENTROPIQUE APPLIQUÉE AUX SYSTÈMES COMPLEXES

1. Complexité : le système doit être suffisamment complexe ; il doit posséder un nombre suffisant de degrés de liberté pour que l'entropie thermodynamique puisse s'y appliquer, et présenter une capacité de changer, une certaine adaptabilité selon les contraintes de l'environnement.
2. Interactions : un apport d'énergie de l'environnement, une consommation de travail et un rejet d'entropie dans l'environnement ; l'identification de l'entropie est parfois métaphorique et elle est alors associée à l'énergie inutilisable ou aux déchets.
3. Structure : celle-ci doit apparaître dans certaines conditions d'apport d'énergie, mais sa caractérisation est largement imprécisée ; on y retrouve les notions de forme, de configuration, de structure ordonnée, d'ordonnancement, de motif, de design, etc. (l'exemple paradigmatique étant les cellules de Rayleigh-Bénard).
4. Coexistence : la stratégie identifie des processus de transformation d'énergie, de travail et d'entropie en concomitance avec l'apparition d'une structure particulière au sein du système ; la question est de savoir quel lien entretiennent ces transformations et cette apparition : un lien d'antécédence, de causalité, de détermination nomologique ; les uns peuvent-elles expliquer les autres ?
5. Émergence : l'organisation structurale considérée nouvelle exemplifie-t-elle l'émergence, et si oui, en quel sens ?

Ceci dit, l'identification d'une condition ou d'une concomitance ne constitue pas nécessairement une explication. L'air est sans doute une condition nécessaire à la combustion mais elle ne l'explique

pas nécessairement. En outre, le statut de condition (nécessaire ou suffisante ?) attribuée au rejet d'entropie dans l'environnement afin de respecter le principe entropique n'est pas justifié (du moins pour l'instant) car ce qui est identifié se résume (pour l'instant) à une concomitance, à une covariation. Par exemple, Prigogine (1968 : 102) affirme que « le fait que leur organisation, qui généralement s'accroît durant cette évolution, correspond à la décroissance de l'entropie » constitue un trait marquant des organismes vivants.

Comme c'est souvent le cas avec ce genre de projet, et ce semble particulièrement le cas avec l'entropie, plusieurs propositions parfois contradictoires ont été offertes pour expliquer cette apparition ou cette émergence de structures dans les systèmes loin de l'équilibre. (Rappelons le cas de la mesure quantique qui est parfois décrite comme impliquant une augmentation l'entropie du système mesuré et parfois une diminution.) Pour Schneider & Kay (1994), « l'émergence de structures cohérentes s'auto-organisant est la réponse attendue des systèmes tentant de résister et de dissiper les gradients qui les poussent loin de l'équilibre ». De sorte que les systèmes complexes comme des écosystèmes se développent de manière à dissiper ou dégrader les gradients d'énergie de manière plus efficiente (Margulis & Sagan 2000). Cette dissipation d'entropie se fait typiquement dans l'environnement du système complexe. Selon une hiérarchie de niveaux d'organisation donnée, cette dissipation se produit dans les niveaux supérieurs de cette hiérarchie, et si l'on interprète le concept d'entropie comme une « mesure du désordre », alors elle a pour conséquence d'augmenter le désordre dans l'environnement : « Lorsque des systèmes complexes hautement ordonnés émergent, ils se développent et croient aux dépens de l'accroissement du désordre aux niveaux supérieurs de la hiérarchie du système »¹¹.

En revanche, pour Ludovisi, Pandolfi & Tattichi (2005), un écosystème, sous des contraintes données, tend à adopter une configuration qui minimise plutôt la dissipation d'entropie spécifique (définie comme le ratio de la génération d'entropie par la quantité d'entropie incorporée dans la structure du système ou par la quantité d'exergie). Cette mesure serait une « indicateur de la capacité d'un système dissipatif à se maintenir loin de l'équilibre en se débarrassant de son entropie »¹². Cette mesure paraît plus juste, car il ne s'agit pas seulement de rejeter de l'entropie, mais plutôt de transformer l'énergie entrant dans le système en un travail utile, et ce faisant une quantité d'entropie est rejetée à l'extérieur : si la quantité d'énergie entrant dans le système est fixée, plus la quantité de travail est grande, plus la dissipation d'entropie est faible ; mais si la quantité de travail augmente, alors soit la quantité d'énergie entrant dans le système augmente pour une même dissipation d'entropie, soit cette

¹¹ « *When highly ordered complex systems emerge, they develop and grow at the expense of increasing the disorder at higher levels in the system's hierarchy.* » (Schneider & Kay 1994 : 25)

¹² « [...] *indicator of the ability of a self-organizing dissipative system to maintain itself far from equilibrium by pumping out entropy* » (Ludovisi, Pandolfi & Tattichi 2005 : 33).

quantité d'énergie demeure la même mais la dissipation d'entropie diminue, soit les deux, c'est-à-dire que la première augmente et la seconde diminue. Le comportement des systèmes complexes comme les organismes vivants est donc compatible avec le principe entropique. Cette conclusion préliminaire s'appuie toutefois sur des considérations assez générales des systèmes complexes et en particulier des organismes vivants, qui à l'évidence ne sont pas tous connus¹³.

Je propose dans ce qui suit deux exemples d'application de la stratégie généralisée du principe entropique appliquée aux systèmes complexes. Le premier étant le phénomène physiologique de la photosynthèse, souvent considérée un élément essentiel du maintien de la vie. Le second étant la théorie constructale, une théorie proposée récemment par un thermodynamicien afin de rendre compte du « design dans la nature », en prenant toutefois la précaution d'éviter toute référence à l'intentionnalité. Mais avant, je me permets un petit aparté afin de mettre en perspective des résultats antérieurs (Chapitre 5), sur les concepts de probabilité et d'entropie, afin d'éclairer la discussion qui va suivre.

6.3.1.2 PROBABILITÉS, INTERVENTIONS ET ENTROPIE

En définissant l'occurrence d'un phénomène particulier comme très probable, généralement on ne demande pas d'explication particulière au-delà des raisons définissant cette probabilité, car elle définit déjà l'expectative qu'une explication est censée fournir. Par exemple, une explication de l'occurrence du tirage d'une bille noire parmi un ensemble de dix mille billes noires et d'une seule bille blanche revient généralement à *définir* la probabilité de cette occurrence parmi les dix mille et une possibilités : « pourquoi une bille noire a-t-elle été pigée ? », « parce qu'il y avait une chance sur dix mille et une que cela se produise ». Mais si le phénomène est très improbable (comme le tirage de la bille blanche), on exige souvent une explication particulière : soit on attribut le résultat à la chance, ce qui revient à conserver le mystère en réaffirmant que ce résultat était possible malgré tout (mais très improbable), soit on cherche ou désigne, par des hypothèses auxiliaires, la cause de ce résultat. Cette cause est une intervention de l'extérieur d'un genre particulier, comme dans le cas d'une tricherie ou (et cela demande plus de précisions) d'un système ouvert dont l'entropie diminue (indice : il faut observer les billes, les trier et choisir la bonne).

¹³ On estime à près d'un milliard le nombre d'espèces ayant existé depuis l'apparition de la vie sur Terre et que de ce total seulement 1% aujourd'hui existent, et tout juste 20% de ce 1% seraient identifiés par un nom. Et une fraction de ce nombre est assez bien étudiée pour être en mesure de faire des affirmations pertinentes sur leur comportement thermodynamique. Par exemple, de la drosophile, un taxon pourtant très étudié, on connaît environ 800 de ses espèces sur plus de 2 000 (Burian 2005).

Je soutiens ainsi qu'il y a deux façons, deux stratégies globales pour « vaincre les probabilités », c'est-à-dire d'obtenir un résultat qui semble à prime abord improbable. La première, assez évidente, est de produire une multitude d'occurrences parmi la gamme des possibles. Le tirage répété d'une pièce de monnaie en est un exemple. Si la probabilité que la pièce tombe et se maintienne à l'équilibre sur la frange semble pratiquement nulle, elle est en revanche significativement plus élevée au sein d'un très grand ensemble d'essais. Un autre exemple, à mon sens, est celui de la sélection naturelle darwinienne où il y a une production d'une grande quantité de variations servant de matériel à ce processus de sélection (ou plutôt d'élimination). La seconde façon est d'exploiter adéquatement l'information sur un système afin d'y intervenir et de le modifier de manière à obtenir un résultat précis et, toute chose étant égale par ailleurs, improbable. Ainsi, il semble à prime abord hautement improbable que je pige une bille blanche parmi un ensemble de dix mille billes noires. Mais si je les observe toutes une à une ou si l'on m'informe de sa position, en bref si je possède de l'information sur le système, alors la probabilité d'obtenir ce résultat précis, la bille blanche, parmi l'ensemble des résultats possibles, est beaucoup plus élevée. Bien entendu, on pourrait objecter que le protocole expérimental n'est pas le même. C'est précisément ce type d'exemple que reprennent les créationnistes¹⁴ – l'appliquant à tort de manière analogique à l'évolution de la vie – où la probabilité d'obtenir un artefact comme une montre par la conjonction de multiples causes naturelles, sans information et sans être pouvant la traiter et l'exploiter, est pour ainsi dire nulle.

Mais ce dernier aspect, cette seconde façon d'obtenir des résultats *a priori* improbables, exige de plus amples précisions. L'information sur le système n'est pas une condition *sine qua non* à l'occurrence de l'un de ses multiples états possibles, car un état même très improbable demeure par définition un état possible. Il demeure donc possible d'obtenir un résultat improbable *par hasard*. À titre d'exemple, devant la séquence de nombres de un à dix mille, la sélection de n'importe quel nombre, de *ce* résultat particulier, paraîtra également improbable. Ce qui est absurde si l'on croit que l'occurrence d'un résultat improbable exige toujours une explication toute particulière, car il faudrait toujours, à chaque sélection arbitraire, trouver une raison spécifique et toute spéciale à ce résultat singulier. Ce que permet l'information, couplée à une intervention articulée en fonction de celle-ci, est de contourner dans une certaine mesure l'exigence de produire une multitude d'occurrences parmi la gamme des possibles (par ex. un grand nombre de tirages parmi l'ensemble de billes). En d'autres termes, elle permet de gagner du temps. Si une grande séquence d'occurrences parmi la gamme des possibles est produite, si le temps n'est pas un facteur limitant, alors le couple information-intervention n'est pas nécessaire à l'obtention

¹⁴ Ils soutiennent qu'un être intelligent est nécessaire à l'apparition de la vie, du moins d'une vie complexe comme l'être humain.

d'un résultat *a priori* improbable. C'est précisément là où achoppe l'argument créationniste. Car dans le cas de l'évolution, une multitude d'essais, soit une production d'une grande quantité de variations génétiques, est produite sur une longue période, de sorte que l'occurrence apparemment improbable ne l'est plus (à l'instar du tirage de la bille blanche suite à une multitude de tirages, d'essais différents). Après tout, on ne voit pas tout ce qui été éliminé par la sélection.

Ces considérations peuvent être très utiles à la compréhension de la stratégie généralisée du principe entropique. Afin de l'illustrer, prenons l'exemple d'un assemblage de briques : effectuer un travail sur ce système ne change pas son entropie, à moins que la friction y dégage une quantité de chaleur, mais elle implique une augmentation d'entropie dans l'environnement. Si ce travail est exploité en fonction de l'information quant à la position des briques (par ex. en déplaçant une brique qui n'est pas à la bonne place), il est possible d'exploiter cette information afin de placer ces briques selon un certain ordre, plus grand qu'à l'état initial, et ce « degré d'ordre » ou de « complexité » (les définitions sont variées, on l'a vu) peut être associé, du moins on le suppose ici, comme une diminution d'entropie. On peut alors supposer que l'augmentation d'entropie dans l'environnement « surcompense » la diminution d'entropie dans le système, soit la « stratégie généralisée », qui s'offre ici le concours d'une interprétation de l'entropie statistique. Il faut toutefois apporter la nuance suivante, à savoir qu'un travail effectué sur un système ne diminue pas directement l'entropie du système, à moins que ce travail soit utile à son organisation de manière à augmenter la dissipation de chaleur ou de matière hautement entropique dans l'environnement. On ne peut donc pas associer aussi facilement l'ordre spatial à l'ordre de l'espace des phases (Chapitre 5). Il faut donc, aussi, définir correctement ce « degré d'ordre » et ce niveau d'organisation afin de les corrélérer, les cas échéant, avec un changement d'entropie. Ici l'ordre est défini à partir de la correspondance à une règle qui peut (doit ?) faire état de certaines symétries (par ex. une forme pyramidale). Il importe donc de prendre en compte les interactions entre le système et son environnement, autrement une attribution d'une valeur de probabilités peut être complètement biaisée ; à cette enseigne, les cellules de Bénard paraissent hautement improbables, pratiquement impossibles.

En conclusion, l'exigence d'expliquer un phénomène improbable doit tenir compte de trois éléments importants : la façon dont est définie la probabilité et plus particulièrement la gamme de possibles pour un système donné, le temps d'évolution de ce système et plus spécifiquement le nombre d'occurrences parmi cette gamme, et la possibilité d'obtenir de l'information *et* d'effectuer une intervention sur le système. Si la complexité est définie par rapport à un arrangement particulier ou un ensemble spécifique de relation entre sous-systèmes, alors il faut s'attendre à ce que cet arrangement soit aussi spécifique et donc, selon l'équiprobabilité souvent postulée par défaut, peu probable. En ce sens, comme l'affirme Bennett (2003) sans toutefois expliciter son affirmation, la complexité ne peut

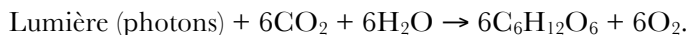
s'accroître rapidement, ou seulement avec une faible probabilité, mais peut s'accroître (de beaucoup) sur des temps très longs (souvent dits géologiques).

6.3.1.3 PHYSIOLOGIE

La physiologie (du grec « *φύσις* », la nature, et « *λογος* », l'étude) étudie le rôle, le fonctionnement et l'organisation physique et biochimique des organismes vivants et de leurs composants. La physiologie étudie (donc) les interactions entre un organisme vivant comme système et son environnement. Elle se prête donc bien à la mise en application de la « stratégie généralisée » du principe entropique selon laquelle, en termes simples, un système ouvert développe une configuration particulière en consommant de l'énergie et en se débarrassant de son entropie. Et comme dans cette étude de systèmes et de sous-systèmes variés, comme la cellule ou le cœur, elle définit schématiquement des niveaux d'organisation, elle paraît ainsi propice à l'analyse du concept d'émergence pouvant y être mobilisé.

Si la physiologie étudie les interactions entre les organismes et leur environnement, elle ne doit pas se limiter à l'influence de l'environnement sur les organismes mais doit aussi s'intéresser à l'influence des organismes sur leur environnement. Turner (2000) propose de se tourner vers la thermodynamique pour une telle analyse car la physiologie à l'intérieur d'un organisme est « fondamentalement un problème de thermodynamique ». Après tout, la thermodynamique décrit les échanges d'énergie sous forme de chaleur et de travail, et les organismes opèrent de telles transformations. Ils le font toutefois en créant ce que Turner appelle de l'« ordonnancement » (« *orderliness* »). On retrouve ici la stratégie généralisée du principe entropique. Mais il faut ici aussi préciser ce que signifie cet « ordonnancement », qui renvoie à l'un des paramètres de cette stratégie avec les notions de configuration, de structure, de motif, de design, d'ordre, etc. Afin de voir si la thermodynamique et la stratégie généralisée du principe entropique peuvent éclairer et même expliquer l'apparition de structures ordonnées, ou leur émergence, tournons-nous vers le processus de la photosynthèse.

Ce processus physiologique consiste en la production de glucose ($6\text{C}_6\text{H}_{12}\text{O}_6$) et d'oxygène (O_2) à partir du dioxyde carbonique (CO_2) de l'air et de l'eau (H_2O) du sol dans un environnement baigné de radiations solaires (lumière). La réaction chimique à l'œuvre dans ce processus peut se résumer ainsi :



L'énergie sous forme de radiations solaires est alors convertie en énergie potentielle emmagasinée dans les liens chimiques de la molécule de glucose que l'organisme peut utiliser plus tard. L'inverse de cette réaction est effectivement ce qui est appelé la respiration, où l'énergie du glucose est consommée par l'organisme pour effectuer un travail physiologique, comme la digestion, et où du gaz carbonique, de l'eau et de la chaleur sont libérés. Mais ce travail ne peut être effectué que si le glucose est couplé à un

processus, appelé cycle ATP (pour adénosine triphosphate), durant lequel l'ATP est d'abord formée à partir de l'adénine diphosphate (ADP) et ensuite, en présence de glucose et d'oxygène, elle libère de l'énergie.

Afin de voir si la stratégie généralisée est respectée, il faut interpréter cette description d'un processus physiologique fondamental et analyser les paramètres de la stratégie en conséquence. L'apport d'énergie ici est sous forme de radiations solaires (lumière) et la dissipation de l'entropie se situe au niveau du rejet d'oxygène ($6O_2$). Comme le bilan énergétique entre les deux ne balance pas, en l'absence d'un autre échange de chaleur ou de l'augmentation de la température interne, il faut qu'il y ait consommation d'une quantité de travail dans le système. Le principe entropique est donc respecté. Concernant la question de la « structure ordonnée », selon l'interprétation du désordre héritée de la version statistique de l'entropie (sans la cautionner ici toutefois), on peut croire que l'augmentation du désordre dans l'environnement résultant de la dissipation de l'entropie entraîne d'une certaine façon une augmentation de l'ordre dans le système. À cette enseigne on joue sur le plan de la métaphore. Turner (2000) soutient que l'ordre ou la « structure ordonnée » – il parle d'« ordonnancement » (« *orderliness* ») – est celle de la molécule de glucose. La raison de cette affirmation repose sur une définition algorithmique de l'ordre (section 6.1), car « moins d'information est requise pour décrire la partie droite de l'équation, et donc elle est plus ordonnée » (Turner 2000 : 13).

Les trois premiers critères de la stratégie généralisée appliquée aux systèmes complexes émergents sont donc satisfaits, nonobstant quelques précisions à apporter. D'abord, les organismes vivants manifestant le phénomène de photosynthèse peuvent légitimement être considérés suffisamment complexes en regard de l'application du principe entropique. Le nombre de molécules et de degrés de liberté est suffisamment grand pour que les concepts d'entropie, de chaleur, de travail et de température aient un sens, et on peut supposer que de tels systèmes opèrent bien au-delà de la transition classique-quantique. Ensuite, il n'est pas besoin de faire appel ici à des métaphores afin d'identifier une certaine quantité d'entropie puisque le rejet d'entropie est directement associé au rejet de chaleur. Enfin, la condition d'apparition d'une structure « nouvelle », ordonnée, est ici celle de la molécule de glucose, qui présente une forme hexagonale (comme les cellules de Bénard). L'analogie ici avec les exemples typiques de la thermodynamique est assez naturelle, car le glucose est assimilable par les cellules vivantes pour lesquelles il est un carburant essentiel. De sorte que la consommation d'énergie impliquant une dissipation d'entropie sert à former une structure ordonnée de manière à produire une capacité de travail, à l'instar d'une transformation énergétique produisant une différence de température (gradient) et donc une capacité à effectuer un travail (Chapitre 4).

La réelle difficulté réside ici dans la caractérisation de cette « nouvelle » structure présentant un certain niveau d'ordre (c'est-à-dire, parmi l'ensemble des propriétés du système, quel sous-ensemble permet de définir cette structure ?). Il y a plusieurs façons d'y arriver. (i) D'abord, le critère de nouveauté est généralement caractérisé en fonction d'un ensemble de lois et de concepts propres à une sous-structure (sous-systèmes ou parties d'un tout) à partir duquel cette structure est jusqu'à un certain point imprédictible, ou à tout le moins il y aurait une certaine hétérogénéité (un terme vague, mais qui a le mérite d'embrasser une diversité de cas) entre cet ensemble et la description de cette structure. Il s'agit d'un obstacle au réductionnisme qui, sans être insurmontable, paraît pour plusieurs auteurs suffisant pour inférer qu'il s'agisse là d'un cas d'émergence (Chapitre 2). En somme, la structure ne serait nouvelle qu'en raison d'une description dont les termes échappent au langage théorique utilisé pour rendre compte de la sous-structure. (ii) Mais ce caractère de nouveauté peut simplement faire référence à la formation diachronique de cette structure. Dit simplement, il y a quelque chose qui n'y était pas. C'est en ce sens que certains (par ex. Bunge 2003) définissent l'émergence selon un critère d'une « propriété qualitativement nouvelle ». (iii) Cette nouveauté ou la structure elle-même peut aussi être caractérisée par le processus duquel elle provient, dont elle en est le résultat (voir la prochaine section). En définissant un tel processus, il est alors possible de définir un lien de causalité entre le processus et cette structure, qui peut être ainsi expliquée selon une théorie causale de l'explication (nonobstant les difficultés que rencontre le concept de causalité). On peut par exemple imaginer un mécanisme de formation d'une telle structure, en l'occurrence la molécule de glucose. Il est aussi possible de restreindre la classe de phénomènes potentiellement émergents par l'apparition d'une propriété structurale (Chapitre 2). (iv) La structure peut aussi être identifiée non par le processus duquel elle suit, mais plutôt des conséquences à la formation de ladite structure, c'est-à-dire qu'une fois la structure en place, le système présente un comportement différent. Par exemple, pour Schneider & Kay (1994 : 31), « l'émergence de la structure ordonnée résulte en une production d'entropie accrue »¹⁵ ; ou encore, la remarquable capacité des organismes d'harnacher l'énergie de la lumière et de l'utiliser pour la synthèse de composés organiques émergeraient de l'organisation structurale dans la cellule (Reece *et al.* 2005 : 186). Ici, tel que mentionné, la capacité d'effectuer un travail conséquemment à la formation de la molécule de glucose servirait de paramètre à cette structure ordonnée.

6.3.1.4 THÉORIE CONSTRUCTALE

La *théorie constructale* est une théorie développée à la fin des années 1990 par Adrian Bejan, un thermodynamicien à l'université Duke, et qui défend la thèse à l'effet que « la génération d'une

¹⁵ « [...] *the emergence of the ordered structure results in more entropy production.* »

configuration de flux est un phénomène physique pouvant être basé sur un principe, la *loi constructale* : pour qu'un système d'écoulement de dimension finie persiste dans le temps (qu'il « vive »), sa configuration doit évoluer de manière à offrir l'accès le plus facile aux courants qui la traversent »¹⁶.

Plusieurs structures observées dans la nature, semble-t-il, proviennent d'un flux de matière, d'un système d'écoulement : le bassin d'une rivière, les vaisseaux sanguins, les branches d'un arbre, la migration d'oiseaux. Y a-t-il une façon de déduire ces configurations, ces formes, avant même de les observer (sans l'aide, selon l'expression maladroite de Bejan, de l'« empirisme ») ? La réponse, selon Bejan, se trouve dans la loi constructale. Le raisonnement est le suivant : tout ce qui bouge, que ce soit animé ou inanimé, est un système d'écoulement (« *flow system* ») ; un système d'écoulement est caractérisé par le courant qui s'écoule (fluide, chaleur, masse, information) et la configuration (« *design* ») à travers laquelle il s'écoule ; si un système d'écoulement possède une capacité, une liberté suffisamment grande de changer sa configuration, alors le système adopte progressivement une configuration, une forme qui facilite cet écoulement. La forme arborescente serait ainsi fréquente dans les formes naturelles car elle favorise de manière efficiente un écoulement d'un point vers une surface et d'une surface vers un point. Dans chacun des cas, la distribution optimale des imperfections est le principe qui génère la forme. Autrement dit, la forme globale « émerge » de la somme d'optimisations locales. Le terme « constructale » provient ainsi de cette idée que la forme globale, la configuration optimale selon les contraintes en présence, n'est pas déterminée *a priori* mais se construit plutôt d'une multitude d'interactions locales qui s'auto-organisent de la plus petite vers la plus grande échelle.

En ce sens, la loi constructale se pose en principe complémentaire aux deux premiers principes de la thermodynamique. Un système présentant initialement des hétérogénéités comme des gradients de température aura tendance à évoluer selon certains flux, en respect des premier et second principes, d'après des échanges avec son environnement et en augmentant son entropie, jusqu'à l'équilibre où il n'y a plus de flux interne. Ou encore, il peut tirer profit de ces échanges afin de diminuer sa propre entropie et ce faisant adopter une configuration ordonnée. Par contre, aucune description de la configuration permettant ces flux n'est offerte, le système étant considéré essentiellement comme une « boîte noire », et c'est pourquoi la thermodynamique est perçue comme une discipline phénoménologique. La théorie constructale dérive les principales caractéristiques de la configuration des systèmes d'écoulement à partir de la maximisation de la performance d'écoulement (« *flow access* ») du système global d'après les contraintes existantes. En adoptant une configuration optimisée relativement aux contraintes locales, ces systèmes évoluent vers une structure générant un minimum

¹⁶ « *For a finite-size flow system to persist in time (to live), its configuration must evolve in such a way that provides easier access to the currents that flow through it* ». (Bejan 2012 : 3 ; aussi Bejan 1998 et 2006 ; Bejan & Lorente 2006)

d'entropie pour une quantité d'énergie absorbée. Elle permet donc d'expliquer l'apparition (la question de l'émergence reste ouverte pour l'instant) de nouvelles structures au sens où il y a organisation structurale et donc instanciation d'une propriété qui n'est pas possédée par les sous-unités d'un système. La complexité que manifeste cette organisation, définie en première approche comme un grand nombre de sous-unités inter-reliées et présentant une fonctionnalité globale, est un résultat, non un objectif.

Les exemples que donnent Bejan sont nombreux. Une rivière, par exemple, est un cas évident d'un système d'écoulement : un élément de rivière est déterminé par l'optimisation de l'aire formée par la distance d'écoulement de la pluie sur la rive et d'un élément de longueur de cette rivière, en fonction d'un taux donné d'écoulement de la masse d'eau de pluie et de la perméabilité du sol. Ces éléments sont ensuite assemblés afin d'obtenir la forme globale. Les résultats concordent bien avec les observations. Il offre aussi une explication du constat bien connu à l'effet que le taux métabolique des animaux (de grande taille) est proportionnel à la masse de l'animal exposant $^{3/4}$. La prédiction de cette corrélation est obtenue à partir de la minimisation (optimisation) de la perte thermique corporelle et la puissance de pompage cardiaque pour un réseau arborescent de vaisseaux sanguins. Le volume de ce réseau est alors déterminé pour une différence de température donnée, de sorte que le taux métabolique, q , équivalent par hypothèse au taux de transfert de chaleur, est $q = (\text{constante}) V \exp^{(3/4)}$. (Pour plus de détails, voir Bejan & Lorente 2006 ; Bejan 2012.)

Cette théorie s'oppose donc à la *théorie fractale*, qui décrit des structures qui sont similaires à diverses échelles, de sorte qu'elles ne changent pas de forme lorsqu'elles sont agrandies ou réduites à une dimension arbitraire. Plutôt que d'offrir une description fractionnant les formes, soit en procédant de l'échelle la plus large jusqu'à une dimension nulle en un nombre infini d'étapes comme le propose la théorie fractale, la théorie constructale propose une description de l'optimisation locale de l'écoulement selon les contraintes, de la plus petite vers la plus grande échelle. Alors que la première prétend que la complexité (encore une fois, les définitions divergent) est « maximisée », la seconde ne considère pas la complexité comme un objectif et s'en tient à l'évolution vers une configuration devant évoluer de manière à offrir l'accès le plus facile aux courants (flux) qui la traversent. Mais surtout, alors que la théorie fractale bénéficie pour ainsi dire de la liberté de l'artiste dans le choix de son algorithme devant reproduire une configuration naturelle (par ex. un arbre), n'étant contrainte que par la répétitivité de la règle itérative, la théorie constructale offre un principe s'accordant avec l'application de principes physiques en fonction de cette optimisation des contraintes. Alors que la première ne dit pas pourquoi une certaine configuration, comme l'arbre, devrait être telle, la seconde offre en revanche un principe, soit la loi constructale. En somme, l'une décrit, l'autre explique (si tant est que le principe soit valide). Il est vrai cependant qu'il n'y a pas, du moins selon l'auteur, de justification de ce principe, autre que les

résultats, la concordance avec les observations. À tout le moins, il n'y a pas de preuve au sens dérivationnel du terme. Mais c'est aussi le cas de la plupart des lois fondamentales.

L'aspect le plus important de cette théorie est toutefois le suivant. Le choix (s'il s'agit bien de cela) de la structure (ou motif) et du processus (ou règle) n'est pas arbitraire. Il a été vu (Chapitre 2) que les émergences inférentielles, faible et computationnelle, ne permettaient pas d'encadrer le choix d'une règle menant au motif prétendument émergent, ni ce motif lui-même. Autrement dit, une fois que ce choix est établi, alors, et alors seulement, il est possible de qualifier le processus d'émergent ou non. Ici, encore une fois, le processus est déterminé par l'optimisation des contraintes locales, selon les principes physiques pertinents à la situation donnée, dans une séquence débutant au sein du plus petit élément d'écoulement vers de plus grands assemblages. En somme, le contexte détermine la règle de l'ordre associé à la structure, soit cette séquence d'optimisation. Le terme de cette séquence détermine aussi l'organisation structurale du système, qui n'a rien à voir avec une quelconque surprise de l'observateur ou de ses capacités à reconnaître une forme quelconque (par ex. un « glisseur », « *glider* »). Selon Bejan (2012 : 10), la complexité est finie et non infinie, et elle « fait partie du design qui émerge ». L'optimisation des contraintes déterminée par le contexte en vue de faciliter l'écoulement mène ainsi au maintien d'une structure permettant cet écoulement par des échanges d'énergie et d'entropie. En somme, le système absorbe de l'énergie et adopte par construction une structure dont l'organisation maximise certains flux, dont une minimisation de la génération d'entropie. Le maintien de cette structure peut être interprété comme la « survie » du système complexe.

6.3.1.5 REMARQUES CONCLUSIVES

La stratégie généralisée du principe entropique appliquée aux systèmes complexes stipule que de tels systèmes sont capables de présenter une structure dissipative, soit d'exploiter un apport d'énergie de basse entropie (haute exergie) afin de se maintenir loin de l'équilibre en se débarrassant d'une certaine quantité d'entropie dans leur environnement, et ce faisant ils développent une « structure ordonnée ». Les organismes ne se nourrissent donc pas de « néguentropie » à la manière d'un chat lapant son lait. Cette structure ordonnée peut recevoir maintes définitions et interprétations, et il n'y a pas de consensus à ce propos. Généralement, elle est définie par rapport à une spécification parmi une gamme de possibles, comme dans le cas de la définition algorithmique. Mais cela revient à la qualifier d'improbable et soulève plus de questions que de réponses. De plus, la formation de cette structure ordonnée n'est pas expliquée si seule la concomitance avec la diminution locale d'entropie est offerte, à moins de s'en remettre à une interprétation de l'entropie comme « mesure du désordre ». C'est ainsi que Prigogine (1968 : 102) souligne certains traits « marquants » des organismes vivants, dont « le fait que leur organisation, qui généralement s'accroît durant cette évolution, correspond à la décroissance

de l'entropie ». Ce qui est plus prometteur est d'offrir un principe encadrant la formation de cette organisation structurale, ce que propose la théorie constructale, d'après un contexte de contraintes locales. Le choix de la règle n'est donc pas arbitraire, comme dans le cas de certaines théories inférentielles de l'émergence et la théorie fractale. Nul besoin, donc, de faire appel à une quelconque « quatrième loi de la thermodynamique » (Kauffman 1993 et 1995 ; Stewart 2003).

6.3.2 Approche analogique de l'entropie comme explanans¹⁷

L'histoire des sciences regorge d'exemples d'emprunts, de récupérations conceptuelles où sont établies des *analogies*, soit des ressemblances formées par l'imagination : on parle ainsi de courant électrique, de flot économique, d'un esprit fonctionnant tel un ordinateur, etc. Outre ces emprunts à des concepts familiers, les analogies en sciences procèdent parfois d'une récupération d'un concept d'une discipline scientifique vers une autre ayant identifié des similitudes dans le comportement des systèmes étudiés. Parfois ces analogies sont taxées d'être purement métaphoriques, vagues ou constituant un obstacle à la compréhension du phénomène qu'elle visait pourtant à éclairer. L'entropie, à certains égards, en est un exemple (Chapitre 3). Mais il arrive aussi qu'elle constitue une base heuristique fructueuse ou possède certaines vertus pédagogiques. C'est le cas une fois de plus avec l'entropie (*ibid.*). Malgré certains atouts que possèdent les analogies, leurs vertus épistémologiques, comme celles de prouver ou d'expliquer, ne sont pas assurées. La reconnaissance des limites de l'analogie, bien qu'avisée, n'empêche pas l'emprunt d'hypothèses posant certains problèmes théoriques de même qu'épistémologiques, tant en physique qu'en biologie : il faut une justification du pourquoi le modèle explique ce qu'il est censé expliquer malgré ces dissimilitudes (une analogie n'est *pas* une identité).

Ainsi, au lieu de dire qu'un concept comme l'entropie s'applique à des systèmes autres que ceux dont traite la théorie qui lui a donné naissance, soit la thermodynamique, comme les systèmes de la biologie, ce que j'ai appelé l'approche substantielle, il arrive qu'un nouveau concept soit élaboré à partir justement de certaines ressemblances entre deux contextes théoriques pourtant différents, ce que j'appelle l'approche analogique. La stratégie explicative à la base de cette approche, en peu de mots, est donc d'abord d'identifier des similitudes dans des descriptions, *A* et *B*, de certains phénomènes que l'on tente d'expliquer et ensuite d'affirmer que *B* se comporte de telle façon *parce que* *A* se comporte aussi de cette façon. Or, un raisonnement analogique peut être bon et mauvais et les discerner exige davantage que la reconnaissance de ressemblances. Par exemple, on peut bien inférer que l'avion, comme l'oiseau,

¹⁷ Cette section est largement inspirée de (Jodoin 2014b).

possède des ailes, et *donc* que l'avion peut voler ; mais il est absurde d'inférer que l'avion, comme l'oiseau, construit des nids. La validité du premier raisonnement et l'absurdité du second, cependant, ne peuvent être reconnues sur la simple base des ressemblances et des dissemblances. Il faut plutôt se référer à ce qui signifie avoir des ailes et faire des nids, bref à des théories scientifiques ou à des connaissances bien établies. Par contre, une fois la validité établie par des raisons extrinsèques à l'analogie, comme avec des théories scientifiques, l'analogie se dissout. En effet, dans l'exemple précédent, une fois que les informations sur le fonctionnement des ailes, leur comportement hydrodynamique, il devient superflu de faire référence à l'analogie.

En effet, une fois que la raison du pourquoi une analogie est considérée comme « bonne » est révélée, l'analogie se dissout généralement puisqu'elle devient plutôt une homologie, comme identité de propriétés ou de structures¹⁸. Par exemple, le vol de l'oiseau est analogue à celui de l'avion mais aussi homologue par rapport à certaines caractéristiques hydrodynamiques, ou encore, une particule et une bactérie, bien que différentes sous bien des aspects, peuvent très bien être homologues selon une description stricte du mouvement brownien. Si des ressemblances peuvent être suffisantes à l'analogie, toute analogie n'est pas explicative car il importe que les éléments similaires supportant l'analogie soient les « bons » (comme dans le cas de l'avion et de l'oiseau). Autrement dit, une analogie est « bonne » dans le cadre d'une explication si elle contient des éléments homologues par rapport à l'explandum (par ex. les ailes et l'hydrodynamique). Et qu'il y ait des similitudes – d'où l'analogie – n'implique pas que la discipline à laquelle on emprunte soit elle-même exempte de problèmes théoriques, lesquels problèmes peuvent être importés du modèle analogue, ni que l'application à un autre contexte théorique soit légitime. C'est pourquoi l'emprunt théorique s'appuyant sur l'analogie requiert une justification autre que la seule ressemblance formelle.

Dans le cadre de la présente discussion, la question est de savoir si des analogues au concept d'entropie dans sa version statistique peuvent expliquer l'émergence de certains phénomènes. Deux exemples présentant la stratégie explicative présentée un peu plus tôt, l'un tiré de la biologie et l'autre de l'économie, seront utilisés. Dans le cas de l'entropie statistique, le principal problème avec cette stratégie est que l'analogie ne peut expliquer le comportement *B* par le comportement analogue de *A*, si *A*, en fait, ne se comporte pas de la façon dont on le suppose. Plus particulièrement, on ne peut expliquer l'accroissement d'une entropie analogue à l'entropie statistique, sur la simple base de cette analogie, si au demeurant l'entropie statistique n'a pas ce caractère de non-décroissance.

¹⁸ Pour une discussion sur les homologies et les analogies dans les explications biologiques, voir Sober (1993) et Ereshefsky (2012). Ce dernier parle des homologies comme nécessairement liées à l'histoire évolutive ; ce n'est pas le sens retenu ici. Pour la physique, Bunge (1967) parle plutôt d'homomorphisme et d'isomorphisme.

6.3.2.1 PROBABILITÉS, INFORMATION ET ENTROPIE

Le principe entropique dans sa version statistique semble se résumer à cette lapalissade voulant que ce qui est plus probable se produit plus souvent. En fait, l'apparent paradoxe entre un principe physique et un principe trivial découle de la définition d'une fonction mathématique selon une gamme de possibles (volume en phase) et l'asymétrie entre les conditions initiales et finales. C'est ce que j'appelle la *trivialisation du principe d'équipartition probable* : si une telle fonction augmente lorsque cette gamme de possibles augmente, et si à l'état initial cette gamme est plus restreinte, car l'état est connu, qu'à l'état final, qui est inconnu, alors cette fonction ne peut qu'augmenter entre ces deux états. Alors, si l'entropie statistique est une fonction proportionnelle au volume en phase en ce sens, alors elle ne peut qu'augmenter, non pas dans le temps de manière monotone, mais entre les états initial et final. L'exemple du dé est ici utile. Si l'état initial est connu, autrement dit si la gamme des possibles est réduite à une seule possibilité, alors l'entropie est nulle, et si, comme c'est le cas en général, l'état final du dé suite à son lancement est inconnu, c'est-à-dire que la gamme des possibles est complète et comporte six possibilités, alors l'entropie statistique, définie selon cette gamme, est plus élevée à l'état final qu'à l'état initial. Cependant, si l'état initial est aussi inconnu que l'état final, c'est-à-dire si la gamme des possibles est la même ou à tout le moins aussi grande pour ces deux états, alors l'entropie statistique est la même pour ces deux états.

Cette version statistique n'est donc pas triviale, mais il faut y apporter plusieurs précisions. D'abord, l'augmentation de l'entropie dans ce cas-ci n'est fonction du temps qu'en raison de l'asymétrie entre les conditions initiales et finales. Il n'y a aucun opérateur d'évolution temporelle, garantissant un accroissement monotone, qui puisse s'y appliquer. En ce sens, il n'y aucune preuve de l'augmentation de l'entropie ou de sa non-décroissance. Voilà un problème de taille pour le réductionnisme (Chapitre 5). De plus, la version statistique du principe entropique n'apporte pas une preuve de l'asymétrie entre les conditions initiales et finales. La raison de cette asymétrie découle plutôt de la sensibilité aux conditions initiales, même pour les systèmes déterministes, ou de l'application de l'hypothèse quant à cette sensibilité. Cette sensibilité veut qu'une petite perturbation ou une petite différence dans les conditions initiales peut mener, même selon une évolution déterministe, à des grandes différences à l'état final. Or l'état initial n'est jamais connu, du moins pour les systèmes macroscopiques, au point de limiter la gamme des possibles (microscopiques) à une seule possibilité. Et c'est ici que l'exemple du dé peut être trompeur. Car même lorsque l'état initial semble être connu puisqu'il est déterminé par une position sur l'une de ses faces (par ex. « 1 »), il y a une multitude de degrés de liberté (par ex. l'angle du dé, son imparfaite symétrie, etc.) qui ne sont pas pris en compte dans cette description simplifiée de l'état initial *mais qui influencent néanmoins le résultat final*. L'état initial n'est pas aussi connu qu'on voudrait

bien le croire. Pareillement pour un système macroscopique, même si celui-ci est décrit très précisément avec des variables macroscopiques comme le volume et la température, et même si son entropie est très faible, il existe un très grand nombre de degrés de liberté possibles, une grande gamme de possibles, dans sa configuration moléculaire compatibles avec ces variables. Avec une évolution déterministe, plus précisément avec le théorème de Liouville, la grandeur de cette gamme de possibles (volume en phase) se maintient dans le temps, mais en général (comme dans le cas du dé) on élargit, plus ou moins arbitrairement, cette gamme. De sorte que sans autre information sur l'évolution du système dans le temps on assigne à l'état final l'ensemble de la gamme des possibles (dans le cas du dé : six possibilités).

Ensuite, malgré l'impasse dans cette tentative d'apporter une preuve de la non-décroissance de l'entropie statistique, celle-ci met en lumière le comportement des systèmes thermodynamiques. Car ce n'est pas qu'il soit très probable que l'entropie augmente, mais plutôt qu'il soit très probable que l'entropie soit dans un état initial évoluant de manière déterministe vers un état final dont l'entropie est plus élevée (Chapitre 5). Ce que réussit à faire le concept statistique d'entropie est de corrélérer un très grand nombre d'états mécaniques (moléculaires), un très grand volume en phase, avec une entropie thermodynamique élevée. Or, l'entropie thermodynamique augmente dans le temps.

Il y a donc une différence fondamentale entre un système thermodynamique et un dé auxquels l'on tente d'attribuer une valeur de probabilité. Dans le premier cas, on *observe* et *mesure* des variables macroscopiques, dont l'entropie, et on constate que l'entropie thermodynamique augmente. Il n'est aucunement question de la structure microscopique de ce système ni de la probabilité que l'entropie augmente. Cependant, cet état macroscopique est compatible avec un très grand nombre de micro-états de l'espace des phases ou plutôt avec un très grand volume en phase. Avec l'hypothèse de l'équiprobabilité des micro-états ou des éléments de volume en phase, on peut dire que l'état d'équilibre a une très grande probabilité synchronique. De plus, si l'on choisit au hasard un micro-état, alors, dans la très grande majorité des cas, il va évoluer, si on le simule ou calcule, vers un micro-état compatible avec une entropie thermodynamique élevée. Dans le second cas, il n'y a pas de telle relation micro-macro, du moins elle n'entre pas en jeu dans l'utilisation ordinaire des dés. L'augmentation de l'entropie combinatoire que l'on pourrait lui attribuer (en supposant que cela ait un sens), découle simplement de l'asymétrie entre l'état initial, supposément connu, et l'état final, supposément inconnu en raison de la sensibilité aux conditions initiales, qui demeurent insuffisamment connues (dans le détail) malgré la spécification de l'état initial.

6.3.2.2 GÉNÉTIQUE DES POPULATIONS

L'objectif de la génétique des populations est de tirer profit des idéalizations que permettraient les grands ensembles, tout en tentant d'éviter les écueils associés aux difficultés calculatoires et interprétatives des statistiques. De manière similaire à la mécanique statistique, il s'agit en quelque sorte d'expliquer, ici aussi, le niveau macroscopique par le niveau microscopique (voir Jodoin 2013e, 2014b et 2014c). Cet exemple est particulièrement pertinent puisqu'une hypothèse fondamentale de cette discipline, ou plus généralement de la théorie des changements micro-évolutionnaires, stipule que « les phénotypes émergent de l'interaction de l'ADN et de l'environnement » (Templeton 2006 : 1). La question ici est de savoir si un analogue au concept d'entropie peut être légitimement mobilisé au sein d'un explanans pour l'explication de l'émergence de certains traits phénotypiques et, pourquoi pas, de tous les traits phénotypiques possibles.

Sella & Hirsch (2005) proposent ainsi un modèle mathématique de la dynamique évolutionniste à partir d'une analogie avec l'appareillage mathématique de la physique statistique. Ce modèle repose sur une stratégie qui consiste essentiellement à identifier les variables de la génétique et de les articuler dans le formalisme de la physique statistique, plus précisément sa mécanique, afin d'élucider certaines relations de la biologie évolutionniste des gènes et des allèles. Ainsi, les variables canoniques d'état de la mécanique statistique que sont les positions et les quantités de mouvement $\{q_i, p_i\}$ sont analogues aux variables génétiques que sont les nucléotides (ou les allèles) $\{A, G, C, T\}$. Sous certaines conditions d'ordre général, il est possible de décrire une dynamique évolutionniste de type markovien menant à une distribution de probabilité stationnaire d'un génotype fixé, c'est-à-dire du génotype du plus récent ancêtre commun. Cette distribution de probabilité serait donc analogue à la distribution de probabilité de Boltzmann-Maxwell en ce sens qu'elle décrit aussi une distribution stationnaire pour un système à l'équilibre. La dynamique évolutionniste, avant l'équilibre ou avant que le génotype de la population atteigne l'état stationnaire, est déterminée par une fonction d'énergie (fonction de Lyapunov¹⁹) que Sella et Hirsch nomme *fonction de la fitness libre* (« *free fitness function* ») :

$$G = \langle \ln(W_m) \rangle - 1/\nu \langle \ln(P) \rangle,$$

où $\langle \rangle$ représente l'espérance mathématique²⁰, W_m est la fitness moyenne, ν est le coefficient de sélection et $\langle \ln(P) \rangle$ correspond à l'entropie S_H de la distribution de probabilité des génotypes. Cette fonction G s'accroît de manière monotone alors que la population évolue vers une distribution stationnaire, à

¹⁹ Une fonction continue décroissante suivant l'évolution d'un état particulier (orbite) du système dynamique (flot).

²⁰ L'espérance mathématique d'une variable aléatoire est l'équivalent en probabilité de la moyenne d'une série statistique en statistiques. C'est une valeur numérique permettant d'évaluer le résultat moyen d'une expérience aléatoire (processus indéterministe ou stochastique).

l'instar du théorème H de Boltzmann où la fonction H (l'inverse de l'entropie) décroît de manière monotone : « La forme de cette fonction d'énergie est précisément celle de l'énergie libre [de Gibbs], et la maximisation de la fitness libre est précisément analogue à la seconde loi de la thermodynamique²¹. » (Sella & Hirsch 2005 : 9545)

Barton & Coe (2009) développent aussi cette analogie entre physique statistique et dynamique génétique évolutionniste en la généralisant à une gamme plus vaste de situations. En délaissant le critère d'une population dont les génotypes sont près de la fixation, les auteurs ajoutent à la fonction de fitness libre G une quantité U qui est une mesure de la diversité génétique ou de l'effet des mutations, telle que

$$G = \langle \ln(W_m) \rangle + \langle U \rangle + 1/v S_H.$$

Les trois termes de cette équation représentent respectivement les effets de la *sélection*, des *mutations* et de la *dérive* (« *drift* »). L'entropie $S_H = -\langle \ln(P[\prod p_k q_k]) \rangle$, où p_k et q_k sont les fréquences des allèles au locus k pour une population se reproduisant sexuellement, décrit la dispersion des micro-états, soit les fréquences des allèles. Les auteurs définissent aussi une entropie $S_a = \ln(\int dp / \prod p_k q_k)$ qui est une fonction des variables macroscopiques, soit la moyenne du trait z , la variance du trait v et la diversité génétique U (qui n'est pas une observable). Elle est proportionnelle, à l'instar de l'entropie de Boltzmann, au « logarithme du nombre de micro-états compatibles avec un état macroscopique donné » et lie par conséquent « les descriptions macroscopiques avec les descriptions microscopiques » (Barton et Coe 2009 : 319). De manière aussi analogue à l'entropie de Boltzmann, les systèmes (biologiques) évolueraient vers un état macroscopique qui est compatible avec plus de micro-états, soit vers une plus grande S_a : « les populations ont tendance à évoluer vers des états susceptibles d'être réalisés de plusieurs façons (soit vers des S_a élevées) »²² (Barton & Coe 2009 : 317).

L'analogie entre la mécanique ou la physique statistique et la biologie évolutionniste (et plus précisément la génétique des populations) peut être fructueuse : les fruits de cette association seront d'ordre heuristique (empirique) mais aussi épistémologique (théorique) si elle permet d'éclairer les mécanismes de l'évolution. L'analogie est toutefois limitée – comme toute analogie d'ailleurs – et les auteurs cités le reconnaissent :

L'analogie avec la thermodynamique est limitée : la fitness moyenne n'est pas conservée de la même façon que l'énergie ; il n'y a pas de contrainte sur l'accroissement de l'entropie, analogue à la seconde loi ; les

²¹ « *The form of this energy function is precisely that of free energy, and the maximization of free fitness is precisely analogous to the second law of thermodynamics.* »

²² « *Populations tend to evolve towards states that can be realised in many ways (i.e., large S_a).* »

populations ne tendent pas vers la même taille lorsqu'elles sont couplées, de la même façon que les systèmes physiques tendent vers la même température ; et il n'y a pas de principe d'équipartition, qui ferait en sorte que les populations s'étaleraient uniformément dans l'espace des fréquences alléliques. Plutôt, l'analogie découle des propriétés générales des processus markoviens, ce qui justifie une approche fondamentalement statistique de l'entropie (Jaynes, 1983; Le Bellac *et al.*, 2004). (Barton & Coe 2009 : 322)²³

La reconnaissance des limites de l'analogie, bien qu'avisée, n'empêche pas nécessairement l'emprunt d'hypothèses posant certains problèmes théoriques de même qu'épistémologiques, tant en physique qu'en biologie : il faut une justification du pourquoi le modèle explique ce qu'il est censé expliquer malgré ces dissimilitudes. L'objectif principal du présent texte est justement de soulever les problèmes que rencontre la génétique des populations dans sa prétention à l'explication des phénomènes biologiques à partir de l'appareillage conceptuel et formel de la mécanique statistique.

Dans le modèle génétique de Barton & Coe (2009), deux « entropies » sont définies : S_H qui est une fonctionnelle de la distribution des micro-états et S_a qui est une fonction de certaines variables macroscopiques. L'entropie S_H représente l'effet de la dérive et implique l'intégration sur l'espace des séquences (fréquences d'allèles)²⁴. Elle exprimerait la tendance des populations, *ceteris paribus*, à adopter des états macroscopiques pouvant être réalisés selon un volume plus grand dans l'espace des séquences, donc vers des S_a plus grandes. La distribution stationnaire de la population idéalisée est déterminée en fonction de la maximisation de l'entropie S_H , conditionnellement à S_a , soit en fonction des contraintes phénotypiques, de sorte que la stratégie explicative est *inversée* par rapport à celle de la mécanique statistique qui se base plutôt sur l'observation de la tendance vers l'équilibre des systèmes macroscopiques et sur la configuration de l'espace des états pour justifier l'augmentation d'entropie. À l'inverse, donc, le modèle de S_a infère la tendance vers des S_a plus grandes à partir de l'hypothèse qu'il est plus probable que le volume dans l'espace des séquences soit plus grand, autrement dit, que la population tend à évoluer vers des macro-états pouvant être réalisés de plusieurs façons par les micro-états. En clair, ce qui est une conclusion en mécanique statistique devient une prémisse en génétique des populations.

²³ « *The analogy with thermodynamics is limited: mean fitness is not conserved in the same way as energy; there is no constraint on the increase in mean fitness, analogous to the second law; populations do not tend towards the same size when coupled together, as physical systems would tend towards the same temperature; and there is no principle of equipartition, which spreads populations out evenly over the space of allele frequencies. Rather, the analogy follows from the general properties of Markov processes, which justifies a fundamentally statistical view of entropy* (Jaynes, 1983; Le Bellac *et al.*, 2004). »

²⁴ Plusieurs travaux ont récupéré la notion d'entropie en génétique des populations. Par exemple, Vladar & Barton (2011) parlent de l'entropie comme d'une mesure de l'information accumulée et maintenue par l'évolution, et Demetrius, Legendre & Harremôes (2009) parlent d'un paramètre statistique caractérisant la robustesse de la population et décrivant l'incertitude dans l'âge de la mère d'un nouveau-né choisi au hasard.

Cette stratégie explicative est d'autant plus problématique que, non seulement, la dérivation de l'entropie statistique est problématique, mais il en va de même de l'argumentation visant à justifier l'augmentation d'entropie à partir de la mécanique classique. En effet, l'hypothèse en mécanique statistique associant probabilité et volume en phase, sur laquelle repose la justification du modèle de S_a , n'est pas aussi évidente qu'on pourrait le croire (Chapitre 5). Donc, s'il n'est pas justifié en mécanique statistique que l'entropie augmente, il n'est pas justifié d'expliquer un phénomène en s'appuyant sur l'augmentation d'une entropie analogue. Mais en supposant que la mécanique statistique soit en mesure de justifier l'augmentation d'entropie, bref le principe entropique, le succès ici du modèle entropique de la génétique des populations dépend de la satisfaction de certains critères mobilisés en mécanique statistique. Car si le modèle est censé fonctionner malgré tout, en dépit de ces critères, alors nul besoin de faire appel à des analogues et au succès de la mécanique statistique en premier lieu.

La justification de l'augmentation d'entropie fait face (je le rappelle) à de nombreux problèmes (Chapitre 5). Elle fait souvent appel à une version quelconque du principe de raison insuffisante, ou, pourrait-on dire, à l'ignorance, dont la stabilité peut aisément être remise en question. Sur cette base, on infère alors l'équiprobabilité des micro-états afin de définir un volume en phase. Or, si cette équiprobabilité peut sembler injustifiée, que ce soit en raison de l'ignorance invoquée ou, encore une fois, en raison d'une certaine corrélation entre différents allèles, celle-ci demeure en général la seule option possible pour inférer quoi que ce soit du micro-état inobservable, hormis bien sûr les contraintes imposées par les variables macroscopiques. L'équiprobabilité est aussi justifiée par l'hypothèse de l'indépendance statistique, comme quoi les différentes directions et vitesses des molécules ne sont pas statistiquement corrélées. C'est pourquoi, sans équiprobabilité, on voit mal comment expliquer le passage d'une distribution à une autre par la transition vers un volume en phase plus grand si ce volume est indéfini.

De plus, l'évolution de la distribution des fréquences d'allèles, c'est-à-dire vers un volume plus grand et une entropie S_a plus grande, ne semble plus justifiée puisque l'évolution temporelle est censée être régie par une dynamique des micro-états, et non par la mesure d'un ensemble, si grand soit-il. En d'autres mots, un système n'évolue pas vers un état parce qu'il y aurait plus d'états compatibles avec ce dernier. Elle ne peut de surcroît s'appuyer sur une notion comme l'ergodicité qui suppose que *tous* les micro-états du système sont visités. Enfin, il n'est pas prouvé que la distribution stationnaire utilisée soit la *seule* distribution stationnaire compatible avec le modèle à l'équilibre.

Ainsi, la stratégie explicative consistant à expliquer le comportement réel en fonction de l'écart qu'il présente par rapport à un état idéal (modélisé) et résultant de l'action d'une force quelconque perd de sa substance en raison du discrédit qui frappe le modèle (de la population idéalisée). Il n'est d'ailleurs

pas évident, comme l'interprétation moderne du TFSN le montre (Okasha 2008), que l'environnement soit un facteur qui puisse être abstrait de l'explication de l'évolution biologique. De fait, aucun génotype n'est inconditionnellement supérieur en fitness à tout autre génotype dans tous les environnements possibles. En outre, la stratégie explicative consistant à expliquer l'évolution d'un état vers un autre par une tendance vers un état d'équilibre perd aussi en force.

En effet, d'une part, l'entropie en mécanique statistique (comme celle de Boltzmann) dispose de très peu d'outils pour une définition claire et non équivoque, et ce même si elle peut s'appuyer sur une théorie comme la thermodynamique et des observations macroscopiques convaincantes, ou, pour le dire autrement, moins équivoques. Malgré la quantité énorme de degrés de liberté au niveau microscopique, l'évolution temporelle de quelques quantités thermodynamiques (macroscopiques) dépend uniquement des valeurs instantanées de ces mêmes quantités.

D'autre part, cette tendance (qui n'est pas observable comme peut l'être la tendance à l'équilibre des systèmes thermodynamiques) découle de l'adoption de la distribution de probabilité normale (gaussienne) centrée sur la moyenne du trait à zéro car, dans une distribution de probabilité normale, la majorité de la population se situe par définition près du centre (généralement à zéro). Ainsi, il y a une quantité S_n , appelée « entropie » par analogie formelle, qui n'est pas observable et qui tire sa définition d'un modèle dont l'idéalisation est discutable mais qui est censée expliquer le comportement macroscopique (évolution de la moyenne et de la variance du trait) de la population. En effet, l'hypothèse d'équiprobabilité est discutable, de façon générale et tout particulièrement pour les systèmes complexes, tout comme l'est l'hypothèse selon laquelle le volume dans l'espace des séquences tend à augmenter.

Donc, non seulement les éléments conceptuels sur lesquels s'appuie la mécanique statistique manquent à la génétique des populations, mais ces éléments sont aussi discutables au sein même de la mécanique statistique. En ce cas, la discussion précédente montre que les éléments disponibles pour justifier le caractère explicatif du modèle deviennent très limités : que ce soit par une cause (Berry 1995), par une loi (Rosenberg 2001) ou par les probabilités (problèmes de l'équiprobabilité et de la non-réversibilité). En clair, le modèle importe les problèmes de la physique en biologie qui ne satisfait pas toujours les mêmes hypothèses. Si l'entropie statistique n'augmente pas, alors l'entropie analogue ne peut augmenter *parce que* l'entropie statistique augmente.

Or, même si le raisonnement analogique tenait, y aurait-il explication des distributions génétiques, ou encore de l'émergence des traits phénotypiques ? Si l'entropie mesure bien le volume en phase des allèles et qu'il y a bien une tendance vers un volume en phase plus grand, alors la réponse à la première question serait affirmative. En ce qui concerne la seconde question, il faut corrélérer ce volume en phase

aux traits phénotypiques, ce qui n'est pas évident, au-delà de l'affirmation que certains de ces traits peuvent être réalisés par plusieurs distributions différentes, contrairement à d'autres.

Néanmoins, une approche comme celle développée par Sella & Hirsh (2005) et Barton & Coe (2009) trouve leur légitimité dans les desiderata d'une science mathématisée et d'une stratégie explicative faisant appel à des mécanismes sous-jacents en complémentarité avec la biologie évolutionniste. Il faut aussi rappeler que la génétique des populations et la génétique quantitative ont connu des succès probants. Mais il faut davantage : les différences dans certains éléments de l'analogie imposent une justification de son emploi comme base d'explication qui se fait d'autant plus pressante que la physique de laquelle elle emprunte peine à justifier certaines de ses hypothèses. Malgré les limites, par définition, d'une analogie, celle-ci ouvre sur des questions d'importance : y a-t-il émergence ou survenance des macro-états sur les micro-états ; est-il justifié de lier volume en phase et probabilité ; jusqu'à quel point l'interaction entre les constituants microscopiques peut-elle être négligée ; la prise en compte de l'environnement est-elle nécessaire, et dans quelle mesure, à l'explication des systèmes biologiques ; sous quelles conditions générales, le cas échéant, une analogie formelle peut-elle être justifiée pour soutenir une explication ?

6.3.2.3 ÉCONOPHYSIQUE

L'éconophysique est un domaine de recherche scientifique interdisciplinaire qui se propose de résoudre des problèmes économiques en appliquant des méthodes et théories développées pour expliquer des phénomènes physiques complexes relevant notamment de la physique statistique. Autrement dit, elle emprunte des modèles à la physique dans le but d'expliquer comment le comportement volatile et largement imprédictibles des marchés financiers émergent d'un large système d'agents humains. Les banques centrales avec des modèles stochastiques dynamiques d'équilibre général ou certaines sociétés d'investissement tentent de tirer parti des avancées de cette discipline. Puisqu'il s'agit d'une approche analogique, il ne faut pas la confondre avec l'économie écologique, développée par Nicholas Georgescu-Roegen (1971 et 1976) et qui tente de prendre en compte la spécificité de phénomènes environnementaux irréductibles à la logique marchande, ou encore la bioéconomie, insistant sur l'origine biologique des processus économiques (Mayumi 2001 ; Vivien 2005). Ces approches tentent d'expliquer ces processus en montrant qu'ils s'appuient sur l'exploitation de sources matérielles et énergétiques de basse entropie ; Georgescu-Roegen a cru à tort y découvrir la « quatrième loi de la thermodynamique.

Richmond, Mimkes & Hutzler (2013), après avoir présenté la forme différentielle de l'entropie thermodynamique, $\delta Q = TdS$ (Chapitre 4), « supposent par analogie que la forme de la différentielle

inexacte de l'argent δM peut être reliée à la forme de la différentielle exacte dS_e via un facteur intégrant λ »²⁵, de sorte que

$$\delta M = \lambda dS_e$$

où S_e est l'*entropie économique*, d'après l'expression de Georgescu-Roegen, et est sans dimension.

En s'appuyant sur la définition de l'entropie macroscopique de Boltzmann, Spânulescu & Gheorghiu considèrent que toute l'*information* d'un marché boursier peut être associée à une particule dans une enceinte, la « chaudière financière » (« *financial boiler* »), avec une quantité de mouvement donnée. Selon eux, dans ce contexte, il serait « suffisamment plausible d'appliquer les mêmes principes, lois et résultats de la thermodynamique, de la physique statistique ou cinético-moléculaire ». Un « micro-état financier » représenterait toutes les informations et décisions matérialisées dans le prix et le volume transigé par unité de temps. Ils proposent alors de définir un « paramètre macroscopique » (un terme vague, certes), P_M , pour la bourse, où le nombre d'arrangements (Ω) compatible avec une distribution donnée (souvent interprétée comme la probabilité d'un macro-état) est remplacée par la probabilité, W_m , de « succès » (une notion qui n'est pas explicitée par les auteurs) d'un micro-état :

$$P_m = k_m \ln W_m$$

où k_m est une constante spécifique un marché boursier. Mais l'analogie commence sérieusement à montrer des signes de faiblesse lorsqu'ils s'appuient sur le lien entre l'entropie thermodynamique (rappel : $dS = \delta Q/T$) et l'entropie statistique pour affirmer que l'inverse du paramètre P_m correspond à la « température du marché boursier ».

Comme dans le cas du modèle précédent de la génétique des populations, cette approche analogique n'est pas nécessairement invalide et il n'est pas exclu que certaines descriptions soient correctes et certaines prédictions avérées. Cependant, un raisonnement analogique peut être bon comme mauvais et distinguer les deux exige davantage qu'une inférence sur des ressemblances : il faut justifier pourquoi deux systèmes semblables mais non identiques devraient se comporter, dans certaines situations, de la même façon. Comme le suggère Forster & Kryukov (2003), peut-être est-ce la méthodologie probabiliste qui porte fruits et non l'analogie physique.

²⁵ « *suppose by analogy that the inexact differential form of money δM may be related to an exact differential form dS_e via an integrating factor λ [...]* »

6.3.2.4 REMARQUES CONCLUSIVES

Le caractère émergent de l'entropie, son irréversibilité ou sa réalisabilité multiple, est transposé dans cette approche aux systèmes censés exemplifier un analogue du concept d'entropie. La stratégie explicative repose donc sur l'hypothèse que tout système exemplifiant le concept d'entropie possède certaines propriétés émergentes (comme l'irréversibilité ou la réalisabilité multiple de certaines propriétés) et que si *tel* système exemplifie un analogue du concept d'entropie, alors *donc, ce* système possède certaines propriétés émergentes, à l'instar de ce qui découle du concept original. Pour que cette stratégie soit effective, il faut d'abord que l'analogie soit justifiée, que les similitudes identifiées dans des contextes différents permettent d'inférer au comportement semblable voire identique du système auquel est appliqué le concept d'entropie analogue. Il faut ensuite, et surtout, que les caractéristiques de l'entropie sur lesquelles on se base pour inférer ce comportement soient bel et bien celles de l'entropie. Mais expliquer l'augmentation d'une entropie analogue à partir de certaines caractéristiques tout aussi analogues sur la base que l'entropie originale augmente, est nul et non avvenu s'il s'agit d'un concept d'entropie ne pouvant exemplifier une non-décroissance (augmentation), comme le concept statistique d'entropie. Il faudrait alors que l'explication fasse valoir des hypothèses auxiliaires comme par exemple celle du grenage grossier, exploitée afin de montrer qu'il y a bien non-décroissance. Or cette tentative de démonstration est déjà discutable et peut difficilement être transposée à d'autres contextes, mais surtout, elle est dérivée de l'observation que l'entropie thermodynamique augmente, un atout qui n'est pas nécessairement partagé par ces autres contextes où l'on veut appliquer un concept analogue.

6.3.3 Discussion & conclusion

En montrant que tous les êtres vivants résultaient d'une évolution où le milieu agit sur des variations différentielles et héréditaires pour éliminer les individus les moins aptes, la théorie darwinienne a su montrer que l'homme n'était pas fondamentalement différent des autres animaux, ni même des autres êtres vivants. Cette loi universelle, ou prétendue telle, unifiait la diversité et l'origine des espèces – ce « mystère des mystères ». Et elle rejetait, assez héroïquement doit-on le préciser, Dieu du paysage explicatif. Mais la vie demeurait un phénomène particulier, elle était toujours spéciale par rapport aux autres phénomènes naturels. En fournissant une loi ou un principe suffisamment général pour embrasser une diversité de phénomènes, tant physiques que biologiques, en supposant qu'ils sont des instances de la catégorie « systèmes d'écoulement », la théorie constructale permet de faire un pas de plus vers cette « désacralisation » : un organisme vivant devient un type de système physique (système

d'écoulement) comme un autre. Nul besoin donc d'un « élan vital ». C'est pourquoi d'aucuns la présentent comme la « loi de la vie » (« *law of life* » ; Basak 2011).

Il s'agit là de l'approche substantielle et de la stratégie explicative généralisée du principe entropique : un système suffisamment complexe en interaction avec son environnement par des échanges d'énergie et d'entropie adopte une configuration particulière qui a comme conséquence un écoulement facilité et une diminution de la génération d'entropie. Les paramètres en jeu sont les suivants. L'entropie est thermodynamique, qui est donc définie en termes de chaleur et de température, mais l'expression du second principe de la thermodynamique, stipulant qu'une quantité de chaleur ne peut être complètement convertie en travail et qu'il y a donc une tendance à la transformation de l'énergie sous forme de chaleur, est aussi pertinent à l'application de cette stratégie explicative. Dans l'exemple physiologique de la photosynthèse, en posant l'hypothèse que la formation de la molécule de glucose nécessite un travail, on doit déduire l'apport d'une quantité d'énergie et le rejet de chaleur et donc d'entropie. Un autre paramètre est l'apparition d'une configuration particulière, qu'on nomme organisation structurale ou structure ordonnée. La caractérisation de cette dernière, bien que souvent ambiguë, est très importante puisqu'elle définit non seulement un explanandum en général, mais aussi le caractère émergent du phénomène, conditionnant ainsi une conception de l'émergence.

La question du pourquoi telle configuration devrait être expliquée parmi l'ensemble des propriétés d'un système revêt alors une importance capitale. Tel que discuté, généralement, un phénomène considéré comme naturel ou très probable ne reçoit pas d'explication au-delà de la réaffirmation du pourquoi il est considéré naturel ou très probable, contrairement à un phénomène bizarre ou improbable. Un phénomène improbable peut être expliqué par l'énonciation des conditions d'attribution de sa valeur de probabilité et d'une longue séquence d'essais ou de répétitions sur une longue période, ou encore par une intervention particulière sur ce système, ou les deux. Ici, les différentes configurations discutées, comme la molécule de glucose ou les cellules de Bénard, présentent un très faible probabilité compte tenu de l'espace des phases initial. Or, il n'est pas question d'expliquer leur occurrence par une longue séquence d'essais mais plutôt par une intervention sur les systèmes qui les manifestent. L'intervention n'est pas ici celle d'un observateur recueillant de l'information de manière prospective à cette intervention. Elle procède plutôt de l'interaction entre le système et son environnement où le premier reçoit une quantité d'énergie et le second une quantité d'entropie. Le problème avec cette interprétation, qui affecte aussi la stratégie généralisée du principe entropique, est ce constat que plusieurs systèmes présentent ce genre d'interaction sans pour autant manifester de comportement complexe ou d'organisation digne d'intérêt.

Dans ce contexte, l'élément important de la théorie constructale consiste à offrir les conditions de réalisation de cette configuration complexe (encore une fois, la diversité des phénomènes et des appellations rend difficile le choix d'un terme générique) selon les conditions locales et l'optimisation ou la facilitation des flux en présence. Cette théorie aspire donc à une très grande généralité car il semble évident que plusieurs entités peuvent être considérées comme l'objet d'un flux et constituant un système d'écoulement, que ce soit l'énergie, la matière ou l'information. C'est pourquoi l'occurrence d'une configuration *a priori* improbable peut recevoir une explication en fonction de l'énonciation de ce processus d'optimisation. Ainsi, ce n'est pas tous les processus d'échange d'énergie et d'entropie qui sont susceptibles de présenter ce genre d'organisation et de complexité, et ce n'est pas non plus tout motif, toute forme ou configuration qui nécessite une explication particulière. De plus, l'optimisation selon les contraintes locales détermine le processus plutôt qu'une règle arbitraire tenant de reproduire une forme, comme c'est le cas de certaines simulations algorithmiques.

Selon le schéma d'analyse conceptuelle proposé, c'est ainsi que l'entropie (en un sens *D*) explique (en un sens *E*) l'émergence (en un sens *F*). C'est-à-dire que l'entropie thermodynamique, au sein de la stratégie explicative généralisée du principe entropique, explique selon un schème causaliste et nomologique, l'émergence de nouvelles configurations, d'organisations structurales. Plus précisément, l'explication revient à circonscrire une trajectoire dans l'espace des phases, dans une gamme de possibles, cette trajectoire étant définie par un principe, le principe entropique, et des conditions locales où plusieurs facteurs, encadrés par des lois, produisent (ce qui peut s'interpréter selon un schème causaliste) la configuration du système. Il s'agit donc d'un type d'émergence essentiellement diachronique, puisque la différence se situe dans le temps, autrement dit le terme d'une trajectoire dans l'espace des phases. Cependant, il y a une relation de dépendance synchronique entre les propriétés des sous-unités et les propriétés structurales l'unité. C'est pourquoi il peut être opportun de définir une *émergence configurationnelle*, où le système adopte une configuration particulière définie par une région spécifique de son espace des phases, qui semble *a priori* improbable, mais dont l'occurrence est déterminée par une interaction particulière avec l'environnement.

La stratégie explicative de l'approche analogique est différente puisque le caractère émergent de l'entropie, son irréversibilité ou sa réalisabilité multiple, est transposé aux systèmes censés exemplifier un analogue du concept d'entropie. Elle repose donc sur l'hypothèse que tout système exemplifiant le concept d'entropie possède certaines propriétés émergentes (comme l'irréversibilité ou la réalisabilité multiple de certaines propriétés) et que si *tel* système exemplifie un analogue du concept d'entropie, alors *donc*, ce système possède certaines propriétés émergentes (à l'instar de ce qui découle du concept original). Cette stratégie ne peut être effective que si l'analogie est justifiée, que les similitudes identifiées dans des contextes différents permettent d'inférer au comportement semblable voire identique du

système auquel est appliqué le concept d'entropie analogue. Or, ce travail de justification constitue souvent une heuristique valable. Car il permet parfois de montrer pourquoi des systèmes différents se comportent néanmoins de manière semblable, donc d'offrir une explication de type unificationniste. Ainsi, dans le cas du modèle de la génétique des populations discuté, il faut dire pourquoi l'association entre volume en phase et probabilité, permettant d'inférer l'équiprobabilité, est justifié malgré les corrélations entre allèles. Mais il faut surtout justifier l'inversion de la stratégie explicative de la mécanique statistique en tenant pour acquise l'augmentation de l'entropie statistique et donc justifier la prémisse génétique qui était une conclusion physique.

Cette approche explique ainsi l'émergence de certaines propriétés à partir bien sûr de propriétés analogues de l'entropie mais aussi à partir d'une association entre structure (ou unité) et sous-structure (sous-unité). Autrement dit, les micro-états sont déjà corrélés avec des macro-états. De sorte que l'explication cible la raison derrière le fait que les systèmes évoluent vers telle ou telle région de l'espace des phases, qui est donc associée à certaines propriétés de la structure macroscopique, que ce soit, dans ce qui a été discuté, certains traits phénotypiques ou d'un paramètre macroscopique boursier. L'explication est donc orientée vers une trajectoire dans l'espace des phases (au sens propre comme au figuré) indiquant pourquoi telle sous-structure (micro-état) est exemplifiée plutôt que telle autre. Maintenant, il n'y a pas d'explication spécifique visant à expliciter ce lien entre structure (ou unité) et sous-structure (sous-unité). Celui-ci est déjà donné par la conjonction des théories où l'une joue le rôle en quelque sorte de la cause formelle, ou plutôt des conditions limites de la dynamique de la sous-structure (micro-état) régie par l'autre théorie. L'explication est donc diachronique mais l'émergence est synchronique. La situation est similaire à celle d'une explication ne disant pas pourquoi tel individu est malade lorsqu'il mange des pommes avariées, mais plutôt pourquoi il mange souvent de telles pommes. C'est pourquoi, par contraste à l'émergence configurationnelle définie précédemment, il est opportun de définir l'*émergence dispositionnelle*, qui internalise justement cette association entre micro-états et macro-états, une sorte d'application (*mapping*) entre deux espaces des phases. Elle se rapproche de l'émergence combinatoire ou méréologique mais s'en distingue puisqu'elle ne se limite pas aux relations méréologiques.

Selon le même schéma d'analyse conceptuelle proposé, il y a donc d'autres sens à donner à l'entropie comme explanans de l'émergence (soit à D , E et F). Le concept d'entropie mobilisé est un analogue d'un concept original, souvent la version statistique, puisqu'il est plus abstrait et se laisse davantage appréhendé dans des contextes différents où des paramètres à un niveau d'organisation circonscrivent le domaine de variation des variables à un niveau inférieur. L'explication est aussi analogique mais de type nomologique, se basant sur le comportement nomique de l'entropie, en l'occurrence sa non-décroissance. Que l'entropie augmente est un fait, mais que la mécanique

statistique soit en mesure d'en apporter une preuve est beaucoup plus problématique. Par conséquent, une explication analogique s'appuyant sur une telle preuve est douteuse. Les outils conceptuels disponibles à la mécanique statistique, qui sont au demeurant problématiques, comme le grenage grossier ou l'ergodicité, sont difficilement « exportables ».

6.4 Conclusion

Les systèmes complexes présentent typiquement plusieurs caractéristiques dont l'analyse, au demeurant difficile par le sens même que l'on peut attribuer à la complexité, gagne beaucoup d'un approfondissement du concept d'entropie. On leur attribue en effet ce caractère hiérarchique où plusieurs niveaux d'organisation sont manifestes et semblent interagir, tandis que le concept d'entropie reçoit au moins deux significations principales, rattachées justement à des niveaux différents. Cette hiérarchie serait d'ailleurs le résultat de processus irréversibles (je n'ai pas traité explicitement cette question toutefois). De plus, alors que les probabilités constituent un outil indispensable à la description de ces systèmes, l'entropie peut à juste titre être considérée comme étant définie à partir d'une mesure de probabilité. Enfin, ces systèmes montrant un haut degré de fonctionnalité requièrent un certain niveau d'organisation et d'ordre, alors même que l'entropie est interprétée comme une mesure du désordre.

Quelle est la raison de cette correspondance si forte entre un ensemble de systèmes différents au comportement générique difficilement identifiable mais néanmoins globalement caractérisé, d'une part, et un concept issu de la physique, d'autre part ? Sans doute que ce concept est un témoin privilégié de ces efforts de conceptualisation et d'explication théoriques alors que nos moyens d'investigation, nos modes d'accès, se précisent et se diversifient. En effet, la motivation de traduire un concept dans un autre langage théorique que celui qui l'a vu naître exige déjà la disponibilité d'un langage alternatif et des modes d'accès qui lui sont propres. La preuve de l'existence des atomes, subséquente à l'entreprise calculatoire de la mécanique statistique, n'a pas modifié sa méthodologie, et à cet égard, s'il y a pu y avoir un changement de perception quant à ses fondements, il semble s'être opéré sur le plan métaphysique et non épistémologique. Dès lors, le statut épistémologique relatif de la mécanique statistique par rapport à celui de la thermodynamique devient plus flou. On peut bien découvrir de nouvelles vertus à la première en raison de l'affirmation de l'existence des entités dont elle décrit le comportement, mais de là à dire qu'elle puisse vraiment expliquer le compte rendu de nos observations empiriques est une autre histoire. La stratégie proposée par Wimsatt (2012), comme quoi l'on doit expliquer avec ce qui est robuste ce qui l'est moins, ne peut donc pas être appliquée aussi clairement qu'on le voudrait. Son critère de robustesse stipule en effet qu'une entité est robuste si elle est accessible d'une multitude de manières indépendantes. On se demande alors si, assez curieusement, la polysémie du concept d'entropie ne deviendrait pas un atout si ces multiples définitions issues de différentes dérivations constituaient justement des modes d'accès indépendants. Malgré cela, cette polysémie et les différentes hypothèses à la base de ces dérivations rendent la tâche particulièrement difficile quant à la question des relations inter-théoriques, dont celle exprimée par l'émergence.

En effet, si l'irréversibilité thermodynamique est la propriété essentielle de l'entropie dont il faut rendre compte, le choix de la base d'émergence est plus délicat. Entre les propriétés mécaniques des molécules et celle d'irréversibilité macroscopique se glissent de nombreuses entités et concepts. L'irréversibilité thermodynamique serait-elle émergente parce qu'elle est (i) absente de la mécanique ou plutôt parce que (ii) l'entropie statistique ne peut la reproduire ? Selon l'option (i), peu importe le concept d'entropie que la mécanique statistique peut produire, cette discontinuité conceptuelle, au sens d'une propriété qualitative nouvelle, va demeurer et l'inférence au caractère émergent paraîtra justifiée. S'il s'agit de dériver un concept à partir des propriétés des entités, les molécules, alors l'approche de Boltzmann est à privilégier plutôt que la méthodologie des ensembles (infinis et fictifs) de Gibbs. L'approche de Gibbs peut trouver une légitimité toutefois au sein de la conception sémantique de l'émergentisme où il est question de relation entre théories et non entre entités. Dans ce cas, on glisse vers l'option (ii), et il faut se poser la question de ce qui est valable comme « importation » d'hypothèses et de concepts pour être en mesure de reproduire et éventuellement réduire le concept d'une théorie à une autre. A-t-on vraiment réduit l'entropie thermodynamique à la mécanique si le concept emprunte autant à la mécanique qu'aux probabilités ? Il y a, de plus, un ensemble d'hypothèses particulières visant principalement à reproduire l'irréversibilité, qui prend alors la forme d'une fonction en phase non-décroissante, et plusieurs interprétations différentes pouvant présenter cette discontinuité conceptuelle comme indice, sans plus, de l'émergence. Identifier un concept d'entropie comme explanandum de l'émergence, c'est donc choisir un ensemble de propriétés comme candidats potentiels de relata d'émergence ainsi qu'un ensemble d'hypothèses auxiliaires, mais c'est aussi devoir choisir entre plusieurs interprétations.

En quel sens alors, et dans un premier temps, l'entropie (au sens C) peut-elle être l'explanandum (au sens B) de l'émergence (au sens A) ? D'abord, l'entropie peut être l'explanandum de l'émergence au sens où elle en est un cas particulier, qu'elle en instancie le concept. La définition d'un concept adéquat d'émergence, respectant le critère de non-trivialité, permet de définir trivialement une classe de référents pouvant s'y appliquer et une classe de référents ne le pouvant pas. Il y a là une explication au sens faible puisque l'espace des phases est réduit et une explication plus précise demeure possible et même souhaitée. Il s'agit toutefois d'un cadre explicatif assez large n'étant pas spécifique à un concept particulier d'émergence. Par exemple, selon la conception ontologique de l'émergence comme occurrence d'une propriété nouvelle qualitative, cette propriété est située dans un contexte méréologique caractérisé par cette nouveauté. Maintenant, il va de soi que l'on est en droit de demander pourquoi y a-t-il occurrence de telle propriété. L'identification de ce qu'on pourrait appeler une coprésence conditionnalisée pour des niveaux différents situe *per vim definitionis* mais appelle aussi un travail explicatif plus approfondi.

Ensuite, une explication de l'entropie en un sens mitoyen ou intermédiaire est aussi possible, où l'émergence est entendue dans son acception méthodologique. L'émergence méthodologique encourage la recherche d'explication réductionniste autant que possible puisque, tout chose étant égale par ailleurs, la simplicité est préférable à la complexité. De plus, le lien entre unité (ce qui est constitué) et sous-unités (constituants), typique mais non spécifique à une réduction, représente un lien modal pouvant soutenir une explication. Sans statuer précisément sur les conditions structurelles que doit rencontrer une réduction, elle n'infère toutefois à l'explicabilité voire au mystère du phénomène ne pouvant être réduit à un certain niveau d'organisation. Ainsi, un phénomène émergent peut en principe être expliqué, même par une explication d'un certain type réductif (non-dérivationnel). Face à cette diversité de niveaux d'organisation et de modes d'accès, l'émergence méthodologique suspend son jugement lorsque plusieurs modes d'accès définissent des niveaux d'organisation différents sans que leur valeur épistémologique respective ne soit clairement établie. Car, je le répète, une description peut être jugée supérieure pour des raisons représentationnelles même si les conditions d'une explication réductive sont remplies.

Enfin, il est possible d'offrir une explication au sens fort, en ce sens où est défini un élément de la structure effective des processus naturels, ou un processus réel, produisant des phénomènes émergents comme l'entropie. Cela revient à caractériser la relation entre entités à des niveaux différents de manière causale ou du moins selon un lien modal (plus) fort. Évidemment, cette caractérisation sera très certainement une catégorie métaphysique. Bien qu'il s'agisse de l'une des raisons pourquoi un cas d'émergence peut aussi être expliqué, le gain explicatif n'est pas assuré. Car une telle explication au sens fort n'est effective que si la catégorie métaphysique est valide. Autrement, cela revient à stipuler les conditions sous lesquelles un phénomène doit être considéré « émergent », comme dans le cas de l'explication au sens faible (vue précédemment), mais en lui donnant un enrobage ontologique.

En tant que concept relationnel et compromis, l'émergence permet d'éviter certains extrêmes que sa motivation initiale souhaitait surpasser. Tout ramener à la mécanique (ou à un ensemble de lois déterministes appliquées à des entités immuables) ou à la thermodynamique (ou à ce qui est directement observable) semble apporter plus de problèmes que de réponses. En revanche, instaurer une barrière étanche entre ces théories et les niveaux d'organisation qu'elles sont censées décrire n'est guère plus convaincant. Le cadre explicatif de l'entropie comme explanandum de l'émergence qui vient d'être présenté n'opte pas pour un seul concept d'émergence, mais exclut toutefois l'émergence définie comme non-explicabilité. Tel que mentionné, la seule explication qui est exclue est l'explication de type réductif au sens dérivationnel du terme, ce qui correspond à une réduction totale (l'histoire nous dit qu'il s'agit d'un cas rare). Cela est pour le moins dommage car une telle réduction permettrait, en

principe du moins, d'appliquer ce concept réductif à tous les niveaux supérieurs pour une stratégie explicative amplement généralisée.

En ce qui concerne, dans un second temps, le cas de l'entropie comme explanans de l'émergence, il faut distinguer l'approche substantielle de l'approche analogique. Dans le premier cas, l'entropie renvoie à une propriété macroscopique robuste et autonome pouvant être mobilisée au sein d'un explanans de l'émergence de nouvelles structures, tandis que dans le second cas, certaines propriétés de l'entropie (sa non-décroissance) servent d'explanans en étant attribuées à des systèmes en fonction de ressemblances (formelles) avec les systèmes auxquelles l'entropie originale s'applique. L'analyse de ces deux stratégies explicatives a permis de distinguer l'émergence configurationnelle (diachronique) et l'émergence dispositionnelle (synchronique).

La stratégie généralisée du principe entropique stipule qu'un système suffisamment complexe en interaction avec son environnement par des échanges d'énergie et d'entropie adopte une configuration particulière qui a comme conséquence un écoulement facilité et une diminution de la génération d'entropie. En ce sens, le système exploite les ressources de son environnement et développe une configuration à ses dépens en y dissipant de l'entropie. En effet, l'entropie dans ce contexte peut être interprétée comme de l'énergie dégradée, comme un déchet. La configuration que le système développe est une organisation structurale en ce sens où les sous-unités entretiennent des relations mutuelles selon une fonctionnalité globale, qui, selon la théorie constructale, consiste en l'optimisation des contraintes locales en vue de faciliter l'écoulement des flux en présence. Comme elle adopte une forme, bien souvent reconnaissable comme la forme arborescente, elle peut être décrite comme étant ordonnée, c'est-à-dire selon une règle globale révélant généralement des symétries. C'est pourquoi cette organisation structurale est improbable compte tenu de l'espace des phases initial du système. Une improbabilité synchronique qui ne renvoie pas à un phénomène de rareté mais plutôt à un système en interaction avec son environnement, en sorte que le découpage équiprobable de l'espace des phases ne tient plus. Dire ainsi que le système se nourrit d'information ou de « néguentropie » est trompeur. Car l'information n'est ni recueillie par un quelconque processus physique ni exploitée de manière prospective, et la diminution d'entropie est parfaitement prise en compte par le principe entropique sans que ne soit nécessaire de recourir à ce genre de néologisme.

Généralement, on décrit ces structures dissipatives en disant qu'il y a génération spontanée de motifs ou de formes (Nicolis & Prigogine 1977 ; Kauffman 1993). Mais la covariation ou la concomitance n'est pas une notion explicative. En définissant un processus, soit celui de l'optimisation de la facilitation des flux, l'occurrence de l'organisation structurale peut recevoir une explication causale-nomique, car le processus est encadré globalement par le principe entropique et plusieurs facteurs concourent à cette

occurrence à la manière interventionniste. C'est pourquoi la complexité de la configuration n'est pas un objectif et elle n'est pas maximisée, mais plutôt optimisée selon les contraintes locales. De plus, elle n'est pas déterminée *a posteriori* par une règle algorithmique arbitraire (il doit respecter le principe de conservation de l'énergie) et ne dépend pas de l'éventuel effet de surprise d'un quelconque observateur. Cette configuration nouvelle par rapport aux sous-unités (parties) dépend de celles-ci et de leur organisation, sans toutefois s'y réduire, et entretient avec elles un lien modal (constitutif). Il peut donc y avoir émergence en ce sens précis. Aussi, ce processus est à la base du maintien de la configuration et éventuellement de son évolution. Le maintien de cette structure peut alors être interprété comme la « survie » du système complexe. Qu'il soit physique ou biologique, celui-ci se voit subsumé sous un principe, la loi constructale, et c'est pourquoi le phénomène du vivant perd en quelque sorte son statut spécial par rapport aux phénomènes physiques en général. Nul besoin donc d'un « élan vital » ou d'une « quatrième loi de la thermodynamique ». Par conséquent, selon le schéma d'analyse conceptuelle proposé, l'entropie (en un sens *D*) explique (en un sens *E*) l'émergence (en un sens *F*) : le sens *D* correspond à l'entropie thermodynamique²⁶, le sens *E* est causal-nomique et l'émergence *F* renvoie à une propriété structurale nouvelle mais issue d'un processus.

L'approche analogique est assez différente. La stratégie explicative est de type déductivo-nomologique en ce sens où l'explanans est un principe entropique analogue. L'hypothèse de base est que tout système exemplifiant le concept d'entropie possède certaines propriétés et que si *tel* système exemplifie un analogue du concept d'entropie, alors *donc, ce* système possède certaines propriétés, à l'instar de ce qui découle du concept original. Une analogie donc. Plus souvent qu'autrement, la ressemblance permettant d'inférer cette exemplification du concept est la réalisabilité multiple d'un tout et de ses parties, et cette propriété est celle de la non-décroissance et indirectement celle de l'irréversibilité. Pour que cette stratégie soit effective, il faut d'abord que l'analogie soit justifiée, que les similitudes identifiées dans des contextes différents permettent d'inférer au comportement semblable voire identique du système auquel est appliqué le concept d'entropie analogue. Il faut ensuite, et surtout, que les caractéristiques de l'entropie sur lesquelles on se base pour inférer ce comportement soient bel et bien celles de l'entropie. Mais expliquer l'augmentation d'une entropie analogue à partir de certaines caractéristiques tout aussi analogues sur la base que l'entropie originale augmente est nul et non avenu s'il s'agit d'un concept d'entropie ne pouvant exemplifier une non-décroissance (augmentation), comme le concept statistique de l'entropie. Il faudrait alors que l'explication fasse valoir des hypothèses auxiliaires comme par exemple celle du grenage grossier, exploitée afin de montrer qu'il y a bien non-décroissance. Or cette tentative de démonstration est déjà discutable et peut

²⁶ Il n'est pas exclu toutefois qu'une généralisation d'un concept statistique au non-équilibre soit possible et qu'il puisse s'appliquer à cette stratégie généralisée.

difficilement être transposée à d'autres contextes, mais surtout, elle est dérivée de l'observation que l'entropie thermodynamique augmente, un atout qui n'est pas nécessairement partagé par ces autres contextes où l'on veut appliquer un concept analogue.

7 Conclusions générales

RÉSUMÉ. L'émergence est parfois présentée comme ce qui expliquerait l'entropie alors que l'entropie expliquerait l'émergence. Tels quels ces deux énoncés ne peuvent être tous deux vrais. Face à la polysémie déconcertante des concepts d'émergence, d'entropie et d'explication, je soutiens que cet apparent paradoxe peut être résolu formellement ainsi : l'émergence (en un sens *A*) explique (en un sens *B*) l'entropie (en un sens *C*) et l'entropie (en un sens *D*) explique (en un sens *E*) l'émergence (en un sens *F*). Ce faisant, des positions originales quant à plusieurs débats philosophiques ont été proposées, comme celle sur la nature du temps et l'irréversibilité, sur l'interprétation des probabilités et bien sûr de l'entropie, sur l'explication scientifique, le sens de l'émergence, et sur des pistes de solution d'une définition du vivant.

*Il faut penser deux fois à ce que l'on veut,
parce qu'un jour on le possède.*

Félix Leclerc

7.1 Le paradoxe du chapeau et de la fourmi

La fourmi, qui disparaissait sous un chapeau trop grand pour elle, ne pouvait plus aller voir du pays. En désirant reformuler le concept d'entropie thermodynamique, de le vêtir sous de nouveaux appareils, le concept original a disparu. Du moins, il a drôlement changé. Que cela soit ou non une bonne chose dépend en grande partie des desiderata d'une théorie scientifique, car plusieurs dédaignent une science phénoménologique, des quantités physiques qui demandent la définition de l'environnement du système et une stratégie explicative se basant sur un cas limite jamais instancié. Certains pensent ainsi que la fourmi a un physique ingrat qu'il faut cacher (sauf qu'ici la fourmi ne représente pas le microscopique mais le macroscopique). Le concept a changé car il provient bien sûr d'un autre contexte théorique et ce faisant il a perdu son caractère d'irréversibilité. Avec son chapeau, la fourmi ne peut se déplacer : dans sa version statistique, l'entropie n'augmente pas. Pour explorer le monde, la fourmi doit quitter son chapeau : pour expliquer de nouvelles structures complexes, l'entropie doit laisser de côté sa version statistique. À défaut de porter le chapeau sous son poids dans ses déplacements, peut-être faut-il savoir apprécier la fourmi telle qu'elle est – et le chapeau aussi.

Car l'entropie révèle de très grandes ambitions : expliquer des phénomènes observables comme la tendance à l'équilibre thermique ou encore un principe fondamental comme l'impossibilité du mouvement perpétuel. Mais ce n'est pas tout. Que l'on juge que cette ambition ait été réalisée ou non, elle en entraîne une autre : expliquer un principe qui lui-même explique un principe fondamental dans un contexte théorique dont les anticipations ne reçoivent de l'expérience qu'un support très indirect. Affirmer l'existence des atomes et des molécules ne valide pas la mécanique statistique pour autant. Le chapeau existe, mais fait-il vraiment ? En fait, il semble qu'il prenne toute la place car on le voit partout. En effet, la version statistique du concept d'entropie, malgré sa polysémie, est récupérée dans moult contextes théoriques, de la théorie de l'information à la génétique des populations ; ce qui alimente aussi sa polysémie. La fourmi se sent bien seule parfois. J'ai tenté dans cette thèse, entre autres, de montrer dans quel contexte et selon quelle signification le concept d'entropie pouvait prétendre jouer un rôle explicatif. D'où sa fécondité. On a vu en effet que, si la réponse n'était pas simple, la question avait engendré des discussions dont l'issue improbable portait la valeur de leurs gains. En bref, les résultats sont philosophiquement riches, même si les définitions uniques et exclusives des concepts d'entropie, d'explication, et d'émergence n'ont pas trouvé leur terme unique, mais bien des clarifications importantes.

La polysémie de ces trois concepts est en effet déconcertante. Mais on a aussi vu qu'elle pouvait être féconde si l'on adopte une conception assez large, mais pas trop, de l'explication. Le constat, rappelons-le, qui a motivé en majeure partie cette thèse est celui à l'effet que l'entropie soit *a la fois* explanandum et explanans de l'émergence, c'est-à-dire que l'entropie est souvent considérée comme une propriété émergente, tandis que l'émergence de certaines structures organisées serait le résultat d'une dissipation d'entropie, une conséquence du principe entropique. J'ai soutenu que cet apparent paradoxe peut être résolu de manière formelle de la façon suivante : l'émergence (en un sens *A*) explique (en un sens *B*) l'entropie (en un sens *C*) et l'entropie (en un sens *D*) explique (en un sens *E*) l'émergence (en un sens *F*). Il fallait donc que je précise *A*, *B*, *C*, *D*, *E*, et *F*. Sans simplement réaffirmer les raisons de cette polysémie, j'ai ainsi proposé un cadre explicatif à certains phénomènes complexes en physique et en biologie.

7.2 Complexité et probabilités

S'il est évident que la réalité nous apparaît par niveaux au sens où différentes entités se composent et se présentent selon des régularités plus ou moins stables, des atomes aux sociétés, il est en revanche plus difficile de définir ces niveaux et leurs relations mutuelles. Cette tâche a été entreprise d'après la position explicite du réalisme scientifique, selon lequel le monde existe indépendamment de l'esprit humain et il est connaissable au moins partiellement. Cela n'implique pas que les théories scientifiques soient infaillibles, et cette faillibilité n'implique pas que le monde soit inconnaissable. La possibilité que nos descriptions de la réalité soient vraies n'implique pas non plus qu'il soit possible de reconnaître cette vérité, car dire qu'une théorie est vraie et une autre fausse revient à comparer deux ensembles de justifications, dont l'un paraît meilleur que l'autre, et non pas d'en reconnaître directement la vérité. Nul besoin donc d'endosser la thèse à l'effet que le réalisme ontologique n'a pas le « droit à l'erreur » en décrivant « ce qu'est » la réalité. Un enrichissement mutuel entre sciences et métaphysique demeure donc possible. Cette tâche s'appuie aussi sur le « critère de réalité » en termes de robustesse proposé par Wimsatt (1981 et 1994). Une entité, qu'elle soit concrète ou abstraite, est robuste si elle est accessible – c'est-à-dire détectable, mesurable, dérivable, définissable, prédictible, etc. – d'une multitude de manières indépendantes, par différents « modes d'accès », et si elle est robuste, elle est prétendument réelle.

Un niveau d'organisation est un niveau compositionnel où les choses ou entités sont organisées en relations tout-parties de manière relative, c'est-à-dire qu'un tout peut être la partie d'un autre tout à un autre niveau. Ils sont constitués des classes d'entités présentant ordinairement des tailles et des propriétés dynamiques comparables, interagissant typiquement et principalement entre entités de même classe. Ils présentent globalement un ensemble de phénomènes suffisamment réguliers pour former une « extensivité relativement close » (« *apparent rough closure* »), et non pas un amalgame chaotique de phénomènes inintelligibles. Un niveau d'organisation est une description d'un ensemble de propriétés et de régularités suffisamment stable pour être déterminées de manière robuste selon ce critère. Cette conception des niveaux d'organisation est compatible, sans l'impliquer, avec une approche causaliste (interventionniste ou manipulationniste) des propriétés et n'adopte pas nécessairement un fondationnalisme ontologique où des entités élémentaires à un niveau fondamental possèderaient des propriétés intrinsèques, tandis que des entités de niveau supérieur présenteraient plutôt des propriétés structurales. Si un mode d'accès empirique est relatif et spécifique à un certain niveau, un niveau n'est pas spécifique à un mode d'accès. Conséquemment, un niveau d'organisation peut être considéré comme un « niveau de réalité », une « caractéristique fondamentale, non arbitraire et extrêmement importante de l'architecture ontologique de notre monde naturel » (Wimsatt 1994 :

225). Enfin, un niveau d'organisation est caractérisé par les propriétés des entités qui le constituent et qui se révèlent par des modes d'accès généralement empiriques, mais pas uniquement, à partir desquels est décrit l'ensemble de ces propriétés interprétées par nos théories scientifiques, alors que ces théories sont aussi, en partie, constituées par l'intermédiaire de ces modes d'accès.

Une particularité des systèmes complexes est d'être composés de plusieurs sous-systèmes (ou sous-unités) présentant plusieurs degrés de liberté selon plusieurs niveaux d'organisation. Par exemple, un organisme vivant est constitué de cellules dont le fonctionnement est déterminé par des réactions chimiques mais est aussi influencé par l'ensemble de son métabolisme. Cet aspect compositionnel des systèmes complexes est toutefois insuffisant à leur caractérisation fonctionnelle. Si plusieurs caractéristiques permettent de les identifier partiellement, leurs modalités demeurent largement indéterminées, c'est-à-dire qu'on sait très peu de choses sur les relations (causales ou autres) que telle propriété complexe peut entretenir avec un autre. D'aucuns affirment ainsi que la possibilité de distinguer différents niveaux et de former une hiérarchie découle de la présence de l'irréversibilité d'un processus à un certain niveau, ou que la non-linéarité est source d'émergence et d'imprédictibilité, ou encore que des transitions de phase est corrélative à une brisure de symétrie donnant lieu à la complexité et au chaos. Non seulement la liste n'est pas exhaustive mais les caractéristiques qu'elle contient se recoupent ou s'impliquent mutuellement. L'une des conséquences importantes de ces constats est la nécessité de mettre à contribution plusieurs disciplines dans certaines stratégies explicatives (interdisciplinarité) ou de former de nouveaux concepts de manière conjointe et complémentaire (transdisciplinarité). Les discussions sur les concepts d'entropie et d'émergence ont permis d'offrir un cadre explicatif de ces systèmes complexes.

On constate ainsi que les régularités à des niveaux supérieurs exemplifient la réalisabilité multiple et qu'elles peuvent être robustes, même si, dans certaines circonstances, la dynamique sous-jacente mène à des comportements exceptionnels, improbables. Après tout, ce n'est pas surprenant car les niveaux d'organisation ne sont pas totalement indépendants et peuvent entretenir entre eux des liens modaux puisqu'ils sont parfois, bien entendu, constitutifs les uns des autres. Et si, comme c'est le cas de la dynamique moléculaire, de tels comportements exceptionnels ou improbables sont possibles, alors il faut bien s'attendre à ce qu'ils se produisent en effet. En dévoilant les régularités et les aspects des structures des systèmes complexes, les probabilités jouent un rôle important dans ces situations où la prédictibilité des systèmes est compromise, habituellement en raison de certains effets non-linéaires. On peut par exemple former un ensemble de conditions initiales à partir de l'état supposé du système et simuler son évolution dans le temps avec certaines mesures d'incertitude. Mais les probabilités n'ont pas un rôle à jouer uniquement dans le cas de complexité ni même un rôle réservé aux sciences, elles

sont pour ainsi dire omniprésentes. Leur interprétation est d'un grand intérêt philosophique. Je présente succinctement dans ce qui suit les résultats de la discussion à leur sujet.

Tout d'abord, il faut savoir distinguer clairement le calcul (formel) des probabilités de l'attribution (empirique) de valeurs de probabilités. L'introduction de probabilités dans un raisonnement et à plus fortes raisons dans une théorie est ainsi *stipulée*, puisque les probabilités ne peuvent être ni observées ni être déduites d'énoncés exempts de probabilités. Car il n'y a aucune règle dans la théorie des probabilités statuant sur la procédure à suivre pour l'identification d'événements probabilistes *dans les faits*. Si le calcul des probabilités se place sur le plan formel et logique, c'est-à-dire non factuel, en revanche l'attribution d'une valeur de probabilité aux phénomènes naturels procède de justifications qui ne se limitent pas aux règles logiques et mathématiques, elles dérivent d'« inférences informelles » (Bunge 1967a : 90).

L'aspect objectif des probabilités se manifeste, dans un premier temps, dans la description nomique ou mathématique des systèmes, que ce soit dans les inférences mathématiques ou dans la description de l'évolution temporelle d'un système particulier. Il se manifeste, dans un second temps, dans la mesure empirique, c'est-à-dire, dans ce qui permet de corroborer ou de falsifier un énoncé, d'établir une gamme de possibles ainsi qu'une valeur de probabilité spécifique selon cette gamme. Par contre, au même titre que l'induction ne peut encadrer le choix des paramètres pertinents, ni le calcul ni la mesure ne garantissent l'établissement d'une gamme de possibles ou d'une classe de référence. De plus, une fois établie cette gamme ou cette classe, sa fragmentation selon une mesure de probabilités n'est pas déterminée, même si l'on opte généralement pour l'équiprobabilité, pour une mesure de probabilité uniforme. Par exemple, la gamme des occurrences possibles d'un dé, $\{1, 2, 3, 4, 5, 6\}$, peut être déterminée par inférence d'après les symétries qu'il présente ou selon différentes mesures suite à des tirs répétés, mais l'inférence passant de l'observation de symétries ou de fréquences relatives apparemment stables à la probabilité des occurrences n'est garantie par aucun critère, outre le succès des anticipations de ces valeurs de probabilités. Par conséquent, on ne peut identifier les probabilités ni aux fréquences relatives, ni à l'équiprobabilité des possibles. Demeure donc une part d'arbitraire dans l'attribution de valeurs de probabilité dans les faits, de sorte que le schème dualiste de la théorie de la connaissance, celui de la dichotomie objectif-sujetif est ici irréductible.

Dans le cas de la physique, j'ai discuté de l'emploi des probabilités en mécanique statistique, où le micro-état est « synchroniquement sous-déterminé » parmi l'ensemble des micro-états possibles établis par la mécanique d'après le macro-état déterminé par la mesure, et en mécanique quantique, où le micro-état est « diachroniquement sous-déterminé » parmi l'ensemble des micro-états possibles établis par le formalisme d'après des mesures préalables. La situation n'est pas différente à ce dernier

cas pour le tir d'une pièce de monnaie où l'ensemble des états possibles, macroscopiques cette fois, est déterminé par des mesures préalables. Ces exemples tirés de la physique permettent d'établir une distinction entre probabilité synchronique et probabilité diachronique, ou *probabilité d'être* (« *probability of being* ») et *probabilité de devenir* (« *probability of becoming* ») selon la proposition de Leeds (2003).

Un dé présente six occurrences possibles {1, 2, 3, 4, 5, 6} et le résultat projeté d'un tir particulier est bien sûr compatible avec cette gamme de possibles, mais aussi avec un ensemble de conditions initiales, lesquelles demeurent inconnues et indéterminées par la mesure. Il est tenu pour acquis que si ces conditions initiales étaient déterminées, donc connues, alors la gamme de possibles se réduirait à une seule possibilité car ce système qu'est le dé est considéré être régi par des lois déterministes et objectives. Car même lorsque l'état initial semble être connu puisqu'il est déterminé par une position sur l'une de ses faces (par ex. « 1 »), il y a une multitude de degré de liberté (par ex. l'angle du dé, son imparfaite symétrie, etc.) qui ne sont pas pris en compte dans cette description simplifiée de l'état initial *mais qui influencent néanmoins le résultat final* d'après une évolution déterministe. Ainsi, le hasard s'introduit dans l'indétermination des conditions initiales menant à l'indétermination du résultat final parmi la gamme des occurrences possibles. Or, même un système aussi simple qu'un dé possède une sensibilité aux conditions initiales suffisante pour supporter l'inférence à l'effet que son état après une évolution temporelle, un lancé, est indéterminé au point de lui attribuer la gamme entière des possibles. Cette hypothèse est aussi à l'œuvre dans ce que j'ai appelé la trivialisation du principe d'équipartition probable : si une fonction est proportionnelle à la gamme des possibles, et si à l'état initial cette gamme est plus restreinte, puisque l'état est connu, qu'à l'état final, qui est inconnu, alors cette fonction ne peut qu'augmenter entre ces deux états. Mais on peut douter d'avoir en main une preuve de l'augmentation dans le temps de cette fonction.

Revenons alors aux exemples de la physique. En mécanique statistique, d'une part, la probabilité est fonction de la compatibilité entre la mesure obtenue de l'état, qui est donc accessible et macroscopique, et l'ensemble des micro-états possibles, c'est-à-dire *calculés* à partir des lois de la mécanique classique, mais néanmoins sous-déterminés. D'autre part, en mécanique quantique, la probabilité est fonction de la compatibilité entre la mesure obtenue du micro-état, qui est donc accessible et partiellement macroscopique par son couplage à l'instrument de mesure, et l'ensemble des micro-états possibles, suite à une certaine évolution temporelle, c'est-à-dire *calculés* à partir des lois de la mécanique ondulatoire, qui sous-déterminent aussi le résultat de la mesure. Les probabilités expriment donc le rapport de l'objectivation des résultats de mesures et d'une description nomique du système sous-déterminant, de manière synchronique ou diachronique, l'état du système parmi une gamme de possibles. Par conséquent, la mécanique quantique pourrait être aussi déterministe que la mécanique classique newtonienne *si* l'état initial était sous-déterminé par la mesure, comme l'est le micro-état en mécanique

statistique. Le schème dualiste objectif-sujetif est, ici aussi, irréductible : l'objectivation des résultats de mesure est la même, mais en mécanique la part d'arbitraire s'installe dans l'attribution d'une mesure uniforme au micro-état synchroniquement sous-déterminé, tandis qu'en mécanique quantique elle s'applique à l'identification de la fréquence relative des micro-états diachroniquement sous-déterminés.

Il y a là matière à de plus amples développements. Cette approche des probabilités ne désigne pas de gagnants entre la théorie subjectiviste et la théorie objectiviste, mais propose plutôt deux perdants. En focalisant sur la méthode d'attribution empirique de quantités probabilistes où se conjugue mesure et calcul, elle paraît offrir un terrain d'analyse fécond pour une révision, ou du moins une clarification, de la notion même d'objectivité. Elle rapproche ainsi l'interprétation des probabilités d'une étude méthodologique des théories scientifiques en général, entre la formulation d'hypothèses et leur mise en œuvre.

7.3 Un pluralisme restreint

Les débats concernant la nature et la forme de l'explication scientifique n'ont pas perdu de leur intensité. Puisqu'il n'y a pas de consensus quant à ce que devrait accomplir une explication scientifique, il y a lieu d'être dubitatif quant à une éventuelle résolution de ces débats. J'ai soutenu que l'une des raisons de cet inaboutissement résidait dans dilemme de la censure auquel doit faire face une théorisation de l'explication scientifique : une théorie scientifique ne satisfaisant pas un certain modèle d'explication doit-elle néanmoins être considérée scientifique ? Le défi consiste ainsi à naviguer entre une théorie trop générale, qui risque la trivialité ou l'incompatibilité avec la diversité des explications, et une théorie trop pluraliste voulant en rendre compte de trop près, qui risque alors d'être purement descriptive, sans pouvoir d'unification et sans intérêt proprement philosophique. Car une explication, qu'elle soit scientifique ou non, est censée réaliser plusieurs tâches : montrer que ce qui s'est produit devait se produire ou qu'on était en droit de s'y attendre, faire comprendre, d'unifier plusieurs phénomènes semblables, permettre de prédire, de contrôler, de manipuler.

Bien que cette liste soit loin d'être exhaustive, on constate qu'il y a plusieurs recoupements entre ce que l'on cherche d'une explication et ce que cherche à obtenir le scientifique en général, ce qu'il poursuit ou juge avoir une certaine valeur dans son travail. Avec une portée aussi large, il devient presque consensuel de dire que le « but de la science » consiste à générer des « explications satisfaisantes » (Popper 1991 : 297). Mais identifier ce but défini en ces termes est loin de donner une théorie de l'explication scientifique. C'est lorsque l'on tente de circonscrire plus précisément ce but et de définir plus clairement ce qu'est une « explication satisfaisante » en sciences que les divisions se manifestent. Par exemple, il peut y avoir d'excellentes raisons épistémologiques d'écarter une explication *ad hoc* même si celle-ci peut avoir des vertus pragmatiques importantes, comme celles de permettre de prédire et de contrôler un phénomène particulier. C'est l'une des raisons pour laquelle on peut facilement trouver des écarts considérables entre le modèle et la pratique, exemplifiant ainsi le dilemme de la censure.

Devant les difficultés d'une théorie unifiée de l'explication scientifique, certains ont récusé ce but et avec lui le caractère normatif de l'explication en sciences, celles-ci se limiteraient, selon eux, à décrire, représenter, prédire. Comme si une description, une représentation, ou une prédiction était nécessairement dépourvue de normativité, comme si elle ne pouvait être bonne ou moins bonne, qu'il n'y en aurait pas de moins scientifique que d'autres. Rejeter toute normativité du discours philosophique sur les sciences revient en définitive à le cantonner à des comptes-rendus. Ce caractère normatif apparaît donc important dans le cadre d'un critère de démarcation du discours scientifique

des autres discours plus ou moins rationnels, comme celui du sens commun ou de la poésie, en bref d'un critère de scientificité. Le réputé critère de falsifiabilité (Popper 1935) est sans doute nécessaire mais insuffisant à cet égard, et il n'est plus question de censurer un discours scientifique sur des considérations ontologiques.

Le critère explicatif peut jouer ce rôle normatif et définir le discours scientifique par ce qui est falsifiable *et* ce qui explique est déjà plus restrictif. Il pourrait s'agir d'un autre critère normatif mais celui de l'explication semble plus fort, d'autant plus si l'on endosse la définition poppérienne du but de la science. Malgré les défis de théorisation qu'elle représente, la notion d'explication est centrale en sciences comme dans l'examen critique qu'en propose la philosophie des sciences. En général, une théorie de l'explication scientifique présente idéalement deux classes de pertinence, l'une générale concernant la forme de l'explanans selon les outils épistémologiques privilégiés (causes, lois, unifications, etc.), et l'autre plus spécifique au contexte d'explication où les paramètres devant s'appliquer à l'explanans sont définis (forces, gradients, lois statistiques, etc.). En plus de ces deux classes de pertinence, une explication scientifique doit respecter certaines caractéristiques minimales, soit la cohérence interne, la cohérence externe, l'adéquation empirique ou la testabilité, et l'asymétrie (si *A* explique *B*, alors *B* ne peut expliquer *A*).

Malgré les difficultés de définir une explication, son rôle définit plusieurs autres notions en philosophie des sciences. C'est le cas notamment de plusieurs conceptions de l'émergence. Intuitivement et dans la pratique scientifique, plusieurs objectifs sont ainsi poursuivis avec une explication. Tel que mentionné, il n'est pas évident qu'une théorie de l'explication doive se coller à cette pratique scientifique. Et il n'est pas plus évident que l'atteinte d'un de ces objectifs soit une garantie de la vérité de l'explanans. Face à ce pluralisme d'objectifs poursuivis ou connotés par la notion d'explication, il va de soi que certaines explications peuvent paraître « meilleures » que d'autres en regard du ou des critères adoptés, qui sont eux-mêmes conditionnés par les objectifs poursuivis. Plusieurs explications, généralement dictées par le sens commun, paraissent tout à fait acceptables même si d'autres semblent justement « meilleures », en fournissant par exemple les précisions du discours scientifique. Il y a donc plusieurs stratégies explicatives, c'est-à-dire plusieurs manières de coordonner les efforts mobilisant les outils épistémiques, et parfois même ontologiques, engagés dans l'atteinte de l'un des objectifs associés à une explication scientifique satisfaisante. Cependant, cela ne signifie pas que toutes ces stratégies soient équivalentes et que cette pluralité implique un pluralisme éliminant par le fait même le caractère normatif de l'explication. Tout n'explique pas.

L'orthodoxie de l'explication scientifique renvoie bien sûr, d'abord, au modèle déductivo-nomologique, où il y a subsomption ce qui doit être expliqué, l'explanandum, sous ce qui est censé

expliquer, l'explanans. La difficulté étant d'identifier un énoncé pouvant jouer le rôle d'une loi au sens où il doit être universel et supporter les contrefactuels. Mais on considère généralement comme valables les énoncés qui ne sont pas universels mais exposant une généralisation statistique. Ensuite, l'orthodoxie comporte un autre modèle, complémentaire puisqu'il s'applique à des situations où le modèle déductivo-nomologique ne s'applique pas, soit le modèle causal-mécanique faisant jouer à la causalité le rôle principal dans l'explanans. Pour autant, on ne juge pas les lois comme impertinentes car elles encadrent le choix de processus causal ou de mécanismes pertinents pour l'explication. La stratégie M (pour manipulationniste) proposée par Woodward (2003a) est très influente à l'heure actuelle et stipule, en termes simples, qu'une cause, déterminée par une intervention modifiant la valeur d'une variable, explique l'effet, qui présente aussi un changement de valeur suite à cette intervention. Au cœur de cette orthodoxie se trouve donc la possibilité d'inférer ou de déterminer l'explanandum à partir d'une loi ou d'une cause.

La mise en œuvre de ces stratégies explicatives orthodoxes nécessite toutefois, bien souvent, le recours à des fictions particulières ou à des idéalizations. En effet, les sciences opèrent parfois sur le mode des *comme si*. Pour autant, le discours scientifique n'est pas pur mensonge. Certaines fictions, dites maîtrisées ou simplement idéalizations, sont analogues au cas réel en ce sens où leurs anticipations ne sont pas démenties et peuvent s'approcher graduellement du cas réel, comme un cas limite. Ce qui les caractérise n'est donc pas de nous livrer une description fidèle de la réalité mais plutôt de nous permettre de nous projeter dans la situation qu'elle représente, délivrer un gain cognitif intéressant. Elles peuvent ainsi offrir une description plus simple, faisant l'économie de mots par rapport à une description littérale, ou montrer en quoi des événements différents peuvent mener au même phénomène, que certaines modifications dans la description de l'état initial d'un système conduit à la même prédiction quant à l'état final, bref qu'il y a une tendance exemplifiée par l'idéalisation.

Les critiques des théories orthodoxes – déductivo-nomologique et causal-mécanique – ont surtout porté non pas sur la pertinence de l'approche, mais plutôt sur sa nécessité étant donné les exigences qu'elles placent sur l'explication. C'est pourquoi des stratégies explicatives alternatives seraient non seulement envisageables mais aussi souhaitables en regard de certains objectifs louables poursuivis, intuitivement, avec une explication. En effet, les particularités de l'explication en biologie ont été présentées, de même que les modèles unificationniste et pragmatique. Ce pluralisme ne devrait toutefois pas cautionner un relativisme. Bien que plusieurs objectifs, apparemment antagonistes, puissent être poursuivis avec une explication, tout n'explique pas pour autant. Il convient donc d'identifier clairement les éléments constitutifs de ce qui peut être légitimement considéré comme une bonne explication. En général, si une prétendue explication fournit une « expectabilité » suffisante, c'est-à-dire si l'explanans fait en sorte que l'on soit en droit de s'attendre à l'occurrence du phénomène

décrit par l'explanandum, alors on juge l'explication satisfaisante. J'ai ainsi proposé un modèle d'explication caractérisant l'expectabilité d'après les notions de modalité et de naturalité, et en empruntant à cette diversité des stratégies explicatives.

Dans le modèle déductivo-nomologique, l'expectabilité est satisfaite par la déduction de l'explanandum à partir de l'explanans. Mais dans certains cas, il semble que l'expectabilité soit satisfaite sans qu'il n'y ait explication. Bien qu'une approche causaliste puisse parfois répondre à ce genre d'objections, il importe de spécifier l'expectabilité. Pour ce faire, il faut reconnaître qu'expliquer un phénomène (ou un ensemble de phénomènes) revient à dire pourquoi il est possible avant de dire pourquoi il s'est produit. J'ai donc proposé un modèle d'explication basé sur l'identification d'une gamme de possibles, sur la définition d'un espace des phases pour l'explanandum. En identifiant une région de l'espace des phases, sont définies deux classes d'énoncés contrastives circonscrivant ce qui est possible et ce qui ne l'est pas. Cela permet une explication au sens faible du terme, en ce sens où la région de l'espace des phases n'est pas complètement spécifiée comme dans un cas de déterminisme absolu, mais permet tout de même de circonscrire un ensemble de prédicats pouvant s'appliquer à l'explanandum. Pour une explication plus forte, il faut l'appuyer par un lien modal entre l'explanans et l'explanandum. Les critères pertinents de l'orthodoxie de l'explication, comme les lois ou les causes, permettent évidemment d'établir des liens modaux. Mais, encore une fois, les exigences placées par l'orthodoxie sont trop fortes et écartent plusieurs explications pertinentes ou jugées valides. De plus, définir la causalité comme un ensemble de conditions nécessaires et suffisantes, de manière à déterminer une trajectoire précise dans l'espace des phases, fait face à des difficultés qui minent la pertinence même d'une cause dans l'explication. Enfin, la modalité n'a pas besoin de signifier une nécessité au sens fort, logique du terme. Par conséquent, une explication satisfaisante n'a pas besoin de déterminer une trajectoire précise dans l'espace des phases, seulement de circonscrire suffisamment les régions pertinentes à cette explication.

Pour circonscrire ces régions, une partie de la gamme des possibles, le critère de naturalité est utile. La naturalité est un terme plutôt vague, mais il embrasse une diversité de cas où l'expectabilité est exemplifiée. Elle peut être conférée, par exemple, par la subsomption d'un concept sous une classe de référents similaires ou par association à un « idéal d'un ordre naturel ». C'est pourquoi un cas limite ou un attracteur dynamique, en montrant que plusieurs conditions initiales d'un système évoluent vers un état particulier, peuvent expliquer. Aussi, quoique cela demanderait quelques précisions supplémentaires, dire qu'un événement est probable revient à dire qu'il est naturel qu'il se produise. Enfin, la naturalité peut définir un certain arrière-plan ontologique éliminant les entités impertinentes à l'explication mais qui répondent tout de même à ses desiderata formels.

Bien qu'une discussion encore plus approfondie sur les critères de modalité et de naturalité serait nécessaire, le présent modèle de l'explication est suffisamment clair pour couvrir un ensemble de situations pertinentes à l'émergence et à l'entropie. Par ces deux critères, ce modèle n'est pas purement formel, car, comme le fait remarquer Sklar (2013 : 64), on ne peut tenir pour acquis que la forme de l'explication est suffisante pour faire abstraction de son contenu. Avant de revenir une dernière fois sur ces deux concepts de l'émergence et de l'entropie, quelques mots sur le réductionnisme.

Il a été vu que l'idée de base de la réduction est exprimée par le prédicat relationnel « n'est rien de plus que ». Il s'agit alors en quelque sorte de ramener la diversité à l'unité. Il faut donc spécifier et justifier quels sont les relata, c'est-à-dire à quoi s'applique la réduction, mais aussi la relation elle-même, c'est-à-dire les conditions que les relata pertinents doivent satisfaire afin d'instancier cette relation et donc ce prédicat relationnel. C'est pourquoi j'ai présenté une taxonomie du réductionnisme. Bien que deux grandes approches, épistémologique et ontologique, sont généralement distinguées, celles-ci semblent inexorablement intriquées puisqu'il semble impossible de justifier une assertion réductionniste ontologique sans faire appel à nos moyens épistémiques. La conception orthodoxe d'une réduction épistémologique stipule qu'une théorie se réduit à une autre si et seulement si la première est dérivable ou déductible de la seconde. Le critère de déductibilité est très fort et une réduction totale en ce sens n'est pour ainsi dire jamais réalisée. On obtient plutôt une « image » du concept ou de la théorie que l'on désire réduire. Car les langages des théories réduites et réductives sont généralement hétérogènes, il y a souvent altération de la signification par transfert d'un contexte à l'autre, et bien souvent la déductibilité exige le recours à des lois-ponts ou à des hypothèses auxiliaires qui laisse un doute quant à l'affirmation que le concept ou la théorie prétendument réduit n'est vraiment *rien* d'autre que ce qui est censé réduire. En somme, plus souvent qu'autrement on opère une réduction partielle et non une réduction totale. Dans ce contexte, l'explication peut se référer à un ensemble de concepts et de lois à un certain niveau sans s'y réduire totalement. Ne serait-ce qu'en raison de certains atouts représentationnels ou cognitifs, il peut être avantageux, compte tenu des multiples objectifs de l'explication, de préserver l'entité que l'on cherche à réduire, voire à éliminer.

7.4 Une polysémie ambitieuse et féconde

Le second principe de la thermodynamique est souvent présenté comme l'une des plus importantes lois de la science moderne ou comme l'une des lois que tout individu « scientifiquement informé » devrait savoir (on se rappellera la diatribe de Snow). Il en va de même du principe entropique, censé le recouvrir avec un concept représentant une quantité physique, l'entropie. Or on a vu les prétentions de ce principe. Il est aussi censé expliquer ce principe fondamental de l'impossibilité du mouvement perpétuel de second type, où une quantité de chaleur serait entièrement convertie en travail. Il s'agit là d'une première ambition de taille. Mais il a été lui-même l'objet d'une tentative d'explication dans un tout autre contexte théorie, celui de la mécanique classique. En raison du nombre gigantesque de molécules composant les systèmes typiquement décrits par la thermodynamique, les probabilités et les statistiques ont tôt fait leur entrée dans cette entreprise réductionniste, celle de ramener un concept macroscopique robuste, pouvant être exemplifié par des phénomènes observables, à un concept défini en termes d'entités microscopiques et inobservables. Traduire ainsi un concept aussi fondamental et aussi bien corroboré, étant lui-même à la base d'une traduction de ce principe fondamental qu'est l'impossibilité du mouvement perpétuel, voilà une seconde ambition d'importance. Cette entreprise, au demeurant légitime, a mené à des propositions de définition et d'interprétation variées, qui ont à leur tour laissé place à des récupérations par analogies formelles, parfois discutables, ouvrant ainsi la voie à d'autres définitions et d'autres interprétations. C'est pourquoi le concept d'entropie fait consensus sur deux points : il est fondamental mais personne ne sait vraiment de quoi il s'agit.

Historiquement, ce concept est apparu dans le contexte de la révolution industrielle du XIX^e siècle et fut motivé par un effort technologique d'améliorer le fonctionnement des machines à vapeur, et donc issu des développements de la thermodynamique, avec les travaux de Carnot, Clausius et Thomson, dans un contexte de révolution industrielle. L'apport majeur de Carnot est d'avoir offert une quantification de la conversion de la chaleur, le « calorique », en travail, la « puissance motrice ». Plusieurs éléments de son argumentation font défaut, au premier chef son adoption de la thèse substantielle de la chaleur, soit le calorique, mais le raisonnement global est solide. Il part de l'hypothèse que la production de travail est toujours accompagnée d'un retour à l'équilibre thermique et, à partir de son postulat sur la conservation du calorique, déduit que cette production est due au transport de calorique d'un corps chaud à un corps froid. Suivant un argument de type *reductio ad absurdum* il démontre son théorème à l'effet que les machines réversibles, une idéalisation, présentent un rendement maximal. En corollaire, il établit que la quantité de travail tiré de la chaleur est indépendante des agents mis en œuvre pour la réaliser et que sa quantité est fixée uniquement par les températures des corps entre lesquels se fait, en dernier résultat, le transport du calorique. Ce faisant, il

offrait aussi une expression quantitative simple, dépendant seulement des températures d'opération, de l'impossibilité du mouvement perpétuel, soit l'impossibilité de convertir complètement une quantité de chaleur en travail. Toute l'essence du second principe de la thermodynamique y est.

Kelvin attribuait d'ailleurs la formulation du second principe, sa « proposition II », à Carnot. Selon lui, il est impossible, par le truchement d'agent matériel inanimé, d'obtenir un effet mécanique à partir d'un corps qui serait refroidi à une température inférieure à celle des objets environnants. Il y a ainsi une tendance à l'équilibre thermique où plus aucun transfert spontané, soit sans agent extérieur aux systèmes effectuant le transfert, n'est possible, et donc plus de conversion en travail possible. Kelvin en conclut que la terre deviendra tôt ou tard « inhabitable pour l'homme » (1852, in 1882 : 514), et que cela est « mécaniquement inévitable ». La formulation de Clausius stipule plutôt que la chaleur ne peut passer d'« elle-même » du froid vers le chaud. Elle insiste donc sur une direction privilégiée des processus thermodynamiques, où le monde « spontané » ne requiert qu'une différence de température pour un transfert de chaleur, tandis que le mode « forcé » demande l'investissement d'une quantité de travail.

J'ai proposé de dégager la signification du second principe par une reconstruction théorique basée sur quelques principes et appelée « séquence thermodynamique » : $\Delta T \rightarrow Q \rightarrow W$ [$= (1 - T/T_+)Q = Q_+ - Q_-$]. Ainsi, une différence de température implique un transfert du chaud vers le froid, diminuant ainsi cette différence de température. Et la quantité de travail (W) obtenu d'une quantité de chaleur transférée (Q) est toujours inférieure à cette dernière d'un facteur $(1 - T/T_+)$. Avec cette tendance à la diminution de la capacité à effectuer un travail et à l'uniformisation des températures, il y a une directionnalité aux processus énergétiques des systèmes thermodynamiques puisque l'inverse est possible, c'est-à-dire qu'une quantité de travail peut être complètement convertie en chaleur. La séquence thermodynamique est ainsi compatible avec les formulations de Kelvin et Clausius. Par conséquent, le second principe n'est pertinent qu'à condition que soient bien définis les concepts de chaleur et de travail. Bien qu'il soit généralement admis que ces derniers définissent les deux modes généraux de transferts d'énergie, il est loin d'être évident que le second principe puisse s'appliquer tel quel aux phénomènes microscopiques. Une généralisation à « tous les processus de la nature » ne va donc pas de soi.

Cette directionnalité des processus thermodynamiques est lourde de conséquences, en particulier en ce qui concerne la notion d'irréversibilité. Un détour vers les débats sur la question du temps, toujours actifs, est donc inévitable. Je ne vais toutefois pas revenir sur l'ensemble de la discussion présentée à ce sujet. Deux sens différents doivent être distingués quant à la notion d'irréversibilité. Le premier est la non-invariance par renversement du temps (non-IRT), comme quoi l'ordre de la séquence de ses états,

est similaire dans un sens temporel ou dans l'autre (par exemple si on le fait jouer sur film), de sorte que la séquence des états du processus sont les mêmes mais dans un sens inverse. Qu'une théorie soit IRT signifie qu'elle autorise certains processus et leurs processus inverses. Si la théorie est déterministe, alors elle offre une symétrie de prédictibilité. Mais cela n'implique pas que les processus le soient, car sa description doit être mise en conjonction avec des conditions initiales et le résultat peut être différent suite à une inversion de l'avant et de l'après du processus. Mais l'IRT des lois n'implique pas l'IRT de la séquence des événements du monde et il est toujours possible d'attribuer un ordre et une direction aux séries d'événements, à moins que l'ensemble de l'univers observé retrouve son état initial. Le second sens de l'irréversibilité est l'irrécupérabilité, selon laquelle la transition d'un état initial à un état final obtenu lors d'un processus ne peut être complètement « défait » une fois le processus enclenché. Son application exige la précision de certaines variables, comme le retour à l'état initial d'une partie de l'environnement du système (par ex. celle en interaction), le temps imparti pour ce retour (ce qui renvoie à ce qui est « possible pour nous » et non en principe) et l'équivalence de l'état initial, qui n'a pas nécessairement besoin d'être exactement le même (certaines variables qui ne sont pas thermodynamiques pourraient être jugées impertinentes).

En exprimant l'impossibilité d'une conversion complète de la chaleur en travail et la maximalité des conversions pour des processus réversibles idéalisés, le second principe décrit donc des processus irréversibles au sens de non-IRT et d'irrécupérabilité. On a reproché à ce principe d'être « négatif » et de ne pas exprimer littéralement cette irréversibilité. Or, la stratégie explicative du second principe est d'exprimer un cas limite idéalisé par rapport auquel sont évalués les processus dont la description est susceptible d'être corroborée. En statuant sur ce qui est impossible, ce principe n'en est pas moins fondamental, car il a été maintes fois corroboré, il peut aussi être falsifié et il supporte les contrefactuels. Cela dit, la détermination de ce qui est possible pour un système thermodynamique dépend des conditions de son environnement. Si l'état du système demeure partiellement indéterminé sans référence à cet environnement, il n'en demeure pas moins que le second principe exprime une limite au domaine des possibles d'un système particulier. Aucun indice, à l'heure actuelle, ne laisse croire à la possibilité d'invalider ce principe. Dans une société industrielle comme la nôtre, cette perte irréversible de la capacité d'un système à effectuer un travail a des impacts écologiques importants.

La première ambition du concept d'entropie et du principe entropique est d'exprimer ce principe fondamental. Et c'est à Clausius que l'on doit le concept original de l'entropie. Il a ainsi exprimé par une fonction d'état ces « valeurs d'équivalence » en supposant un cas limite idéal exemplifiant, comme chez Carnot, la réversibilité, où ces valeurs se compensent mutuellement. Cette fonction d'état a été désignée par la lettre S , pour aucune raison apparente, et nommée « entropie » à partir du grec pour « transformation » et du préfixe « en » pour sa proximité avec le mot « énergie ». L'idée générale

derrière sa définition est qu'une certaine quantité, déterminée par le passage d'un système d'un état initial à un état final par un certain processus, ne puisse être compensée par une autre quantité, déterminée cette fois par le passage de ce système de l'état final précédent à l'état initial original par un processus différent. Un processus cyclique est alors défini par une combinaison d'un processus réversible et d'un processus irréversible. Dans le cas limite d'un processus réversible, la fonction d'état est définie telle que $\int dS \equiv \int \delta Q/T = S(x_f) - S(x_i)$ et lorsque ce n'est pas le cas, donc toujours, alors $\int \delta Q/T < S(x_f) - S(x_i)$. L'entropie n'est pas une quantité conservative, et elle est additive et extensive. Elle n'est bien définie que lorsque la relation entre Q et T est déterminée, pour des états d'équilibre, et le résultat n'est pas une valeur absolue d'entropie d'un état mais bien sa différence par rapport à un autre état donné.

Il est évident que l'on puisse remettre en cause l'argumentation faisant état de l'irréversibilité de la théorie, d'après, bien entendu, une interprétation claire de ce qui est signifié par ce terme. Mais la signification de l'irréversibilité est déterminée ici par la théorie qui définit les conditions d'un processus réversible. Si un processus ne remplit pas ces conditions, alors, trivialement, il n'est pas réversible. Et si un processus ne remplit pas ces conditions, alors il possède certaines caractéristiques déterminées par la théorie. Or, encore une fois, la théorie définit les conditions d'un processus réversible comme un cas limite jamais instancié. Ce sont donc les processus qui ne sont pas réversibles qui sont susceptibles d'être observés et leur description corroborée. Une éventuelle critique devrait ainsi porter non pas sur la signification de l'irréversibilité mais plutôt sur cette stratégie explicative visant à définir un cas limite idéalisé par rapport auquel les processus susceptibles de faire l'objet d'une corroboration sont évalués. C'est en ce sens seulement que les processus qui ne sont pas réversibles, donc irréversibles, sont « réels ». Sans cet ingrédient qu'est cette séparation entre un cas limite idéalisé et les autres cas qui sont évalués par rapport à celui-ci, la critique a tout le loisir d'affirmer que la théorie ne traite pas d'irréversibilité.

Le principe entropique stipule que l'entropie ne peut pas décroître. Cela signifie que pour tout système thermodynamique, défini à l'équilibre et plus spécifiquement avec une température déterminée, tout processus ne peut mener qu'à une augmentation de l'entropie totale, incluant celle du système en question et de son environnement. Plus précisément, pour tout processus passant d'un état x_i à un état x_f , le principe entropique veut que $S(x_f) \geq S(x_i) + \int \delta Q/T$, c'est-à-dire que la valeur de l'entropie de l'état final est supérieure ou égale à la valeur de l'entropie de l'état initial plus la valeur de l'intégrale définie par la quantité de chaleur échangée par le système divisée par la température à laquelle s'effectue cet échange. Lorsque le processus est réversible, l'équation précédente devient une égalité telle que $S(x_f) = S(x_i) + \int \delta Q/T$. Il s'agit cependant d'un cas idéalisé jamais instancié. Si le processus est cyclique, ou si le processus est adiabatique ($\delta Q = 0$), alors l'intégrale $\int \delta Q/T$ devient nulle et on a donc

$S(x_j) \geq S(x_i)$. Pour *tout* processus, $dS > 0$ s'applique aux processus irréversibles (et supposément « réels »), $dS = 0$ s'appliquent aux processus réversibles (idéalisés), et $dS < 0$ ne s'applique pas (du moins dans le cas d'un système isolé, ce qui est supposément le cas de l'univers comme un tout). La *génération* ou variation *totale* ou *globale* d'entropie, dS , peut être scindée en deux, soit en une *production* d'entropie ou entropie *interne*, $d_i S$, et un *flux* d'entropie ou entropie *externe*, $d_e S$. Tandis qu'une production d'entropie, qui est un accroissement d'entropie dû aux modifications intérieures du système, est strictement positive, un flux d'entropie, dû aux échanges entre le système et l'extérieur, c'est-à-dire son environnement, peut être négatif pour un système qui est alors ouvert, de sorte que $d_e S < 0$, ce qui correspond à une diminution *locale* d'entropie, c'est-à-dire pour le système en question.

On a vu que le principe entropique était équivalent aux formulations de Kelvin et Clausius du second principe ainsi qu'à la séquence thermodynamique, en ce sens où ils sont violés dans les mêmes circonstances, typiquement thermodynamiques. Mais il demeure possible en principe que des concepts différents, comme la chaleur, le travail, la température et l'entropie, ne soient pas coextensifs et qu'en certaines circonstances bien particulières leur équivalence ne soit pas démontrée. De plus, comme l'entropie est une quantité physique dérivée qui n'est pas directement observable, on retrouve l'objection à l'effet qu'elle n'est définie que pour des cas idéalisés ou que les hypothèses qui supportent sa définition entrave sa capacité à jouer un rôle explicatif important. C'est pourquoi d'éventuelles objections à l'effet que le principe entropique ne renvoie à rien de « réel » ou encore que ce qu'il cherche à expliquer est absent de la définition du concept d'entropie, risquent de s'attaquer *in fine* à une stratégie explicative faisant intervenir des idéalizations. L'une d'entre elles concerne l'équilibre des systèmes thermodynamiques, qui *stricto sensu* n'est jamais atteinte. Mais la condition d'équilibre relève davantage d'une exigence pratique puisqu'il est difficile de déterminer empiriquement et de calculer un état qui n'est pas à l'équilibre. Autrement dit, ce n'est pas que le principe soit violé ou même inapplicable en principe, mais plutôt que certains paramètres soient partiellement indéterminés.

La stratégie explicative du principe entropique est double puisqu'elle explique à partir de causes, soit les irréversibilités, qui sont régies par un principe et qui provoquent cet écart d'un système avec un cas limite idéalisé. Elle est souvent généralisée à des systèmes complexes fort divers où différents transferts de matière et de chaleur sont associés à des transferts d'entropie. Ces derniers sont à leur tour identifiés à une diminution de la capacité du système à effectuer un travail, c'est-à-dire qu'un système qui reçoit une quantité d'entropie (plutôt : certains transferts de matière et de chaleur) subit cette diminution et vice-versa. Bien que le principe entropique dans sa version thermodynamique ne suggère aucune interprétation particulière, on associe souvent l'augmentation d'entropie à une augmentation du « désordre », c'est-à-dire qu'un système qui rejette une certaine quantité d'entropie augmente son niveau d'organisation et vice-versa. Si l'association des transferts de matière et de chaleur à des

transferts d'entropie pour des systèmes complexes et/ou loin de l'équilibre de même que l'interprétation du « désordre » sont justifiées, alors cette stratégie explicative généralisée peut être très puissante puisqu'elle peut rendre compte d'une très vaste gamme de phénomènes. Bien que cette « importation » conceptuelle, de l'hydrodynamique dans le premier cas ou de la mécanique dans le second cas, soit discutable, le principe entropique peut s'appuyer sur la « robustesse » du principe de l'impossibilité du mouvement perpétuel.

La seconde ambition du concept d'entropie et incidemment du principe entropique est de vouloir traduire son expression thermodynamique au sein de la mécanique classique. Ainsi, à la fin du XIX^e siècle, avec les travaux de Maxwell et Boltzmann l'entropie a reçu une nouvelle formulation et donc de nouvelles interprétations. Son traitement mathématique à partir d'une théorie jugée fondamentale, la mécanique, appliqué à un très grand nombre d'entités inobservables, a nécessité l'apport des probabilités et des statistiques. Il est pour le moins ambitieux de vouloir traduire un concept fondamental et bénéficiant du support de l'expérience par l'observation, et sur lequel repose de grandes prétentions explicatives, au sein d'une autre théorie dont la mise en œuvre exige plusieurs hypothèses pour pallier à la sous-détermination empirique des systèmes auxquels elle s'applique. L'opposition entre empirisme et rationalisme trouve donc ici l'une de ses expressions les plus vives. Pourtant, à chaque fois que l'arrangement des sous-systèmes (sous-unités) d'un système (unité) paraît indéterminé, d'aucuns y voient l'occasion d'y appliquer le concept d'entropie statistique.

Les concepts statistiques doivent satisfaire deux propriétés fondamentales qui sont tirées de la formulation thermodynamique puis attribuées à sa formulation statistique, soit la non-décroissance, définie dans le temps pour les processus adiabatiques, et l'additivité, où l'entropie de deux systèmes est la somme de deux entropies individuelles. Ils sont aussi définis selon deux approches, celle de Boltzmann et celle de Gibbs. Dans le premier cas, l'entropie est propriété d'un système individuel définie à partir du théorème H selon lequel toute distribution évoluerait vers une distribution de type maxwellien puisqu'une quantité $H[f_i]$ décroît toujours jusqu'à atteindre un minimum stationnaire dont la distribution est justement de ce type. L'entropie continue de Boltzmann est définie à partir de cette quantité et simule le comportement de l'entropie augmentant dans le temps avant de demeurer constante à l'équilibre ($S_{B,f} \equiv -kNH[f_i]$) ; l'entropie combinatoire de Boltzmann procède à une discrétisation de l'espace des phases ($S_{B,c} \equiv k \ln \Omega$) ; tandis que l'entropie macroscopique de Boltzmann attribue un volume à cette discrétisation. L'entropie de Gibbs, pour sa part, procède par analogie formelle et adopte une stratégie faisant intervenir des ensembles infinis fictifs ($S_{G,f} \equiv -k \int \rho(x) \ln \rho(x) dx$). Dans chacun de ces cas est exemplifié la réalisabilité multiple, qui définit des propriétés de niveau supérieur, un macro-état, étant réalisées de plusieurs façons différentes par des propriétés de niveau inférieur, les micro-états.

Devant cette pléthore de définitions de l'entropie il n'est pas étonnant de constater que plusieurs interprétations ont été proposées. Déjà, les tentatives de la mécanique statistique sont aussi des interprétations à partir d'un autre langage, considérés fondamental et peut-être plus primitif, celui de la mécanique classique. Les propositions sont donc nombreuses et on a vu qu'il n'y a pas de solution facile, pas de réponse à l'emporte-pièce. La fécondité du concept d'entropie est alors manifeste.

Interpréter le concept d'entropie d'après le concept de probabilité découle du projet même de la mécanique statistique. Déjà Boltzmann considérait les « problèmes » de cette discipline comme étant des problèmes du calcul des probabilités (Chapitre 3). Mais devant le nombre impressionnant de discussions consacrées à l'interprétation de ce concept on est en droit d'être dubitatif quant au réel succès d'une telle interprétation comme pouvant favoriser la compréhension. Ce qui détermine les valeurs des paramètres thermodynamiques et macroscopiques détermine aussi le cadre formel dans lequel s'inscrit la gamme des micro-états possibles de la mécanique. Un état macroscopique peut donc être dans l'un ou l'autre de ces micro-états, ou encore il peut être dans telle ou telle région de l'espace des phases. Il semble donc naturel de définir ce rapport entre ce qui détermine l'état macroscopique et ce qui détermine la gamme des possibles microscopiques comme la probabilité de cet état macroscopique. De plus, compte tenu du *fait* que les systèmes thermodynamiques tendent à évoluer vers un état d'équilibre *et* qu'un tel état est typiquement décrit comme un état compatible avec la plus grande gamme de possibles, le plus grand nombre de micro-états possibles ou la plus grande région de l'espace des phases, et de loin, alors il semble aussi naturel d'expliquer cette tendance fortement corroborée comme un processus d'un état improbable à un état plus probable.

Mais cette interprétation est trompeuse. D'abord, on a vu que la méthode d'attribution de valeur de probabilités, qui définit aussi leur interprétation (critère de vérifiabilité), ne pouvait être réduite à l'un des termes du schème objectif-subjectif, à moins de transposer l'observation de faits lors d'une mesure dans le champ de la psychologie du sujet ou d'ériger notre ignorance, par le biais du principe de raison insuffisante, au rang de principe formel justifié. J'ai soutenu qu'on ne pouvait réduire les probabilités aux fréquences relatives ni au découpage symétrique de la gamme des possibles. Par conséquent, toute attribution de valeur de probabilité possède une part d'arbitraire, sans pour autant impliquer qu'aucune attribution de la sorte ne puisse être justifiée. La connexion entre probabilité et entropie est donc subtile. Car le principe entropique ne peut être convenablement appliqué si l'on fait abstraction et de la définition du système (ses frontières) et de son isolation (ses interactions). Si l'on soutient qu'une entropie élevée correspond à une forte probabilité définie par un grand volume en phase, et que, de plus, il est possible que plusieurs systèmes forment une classe dont la majorité est constituée de systèmes de faible entropie, abstraction faite de leurs interactions avec l'environnement, alors ce n'est pas la connexion entre entropie et probabilité qui est jeu, mais plutôt la validité du principe entropique, qui

serait constamment violé. Il s'agit d'une erreur classique, qui a laissé croire que le phénomène du vivant était en contradiction avec ce principe.

L'interprétation du désordre a une source dans certaines intuitions en ceci que l'ordre est davantage associé au travail qu'à la chaleur, mais aussi dans la théorie de la mécanique statistique au sein de laquelle on retrouve la définition microscopique d'un travail comme un transfert d'énergie « de manière cohérente » (Gould & Tobochnik 2010 : 18) ou impliquant « peu de degrés de liberté observables » (Penrose 1970 : 158). Puisque l'ordre se définit par rapport à une règle plutôt que par rapport à des impressions, il faut savoir la justifier. Mais elle est souvent associée à un état spécifique parmi une gamme de possibles (par ex. je pourrais ainsi inventer un mot et croire que la suite de caractères qui le définit comme étant « ordonnée »). En revanche, si un état spécifique est recherché ou si l'on désire le maintenir, des informations et des interventions sur l'unité, dont les sous-unités peuvent prendre plusieurs arrangements, sont nécessaires. Ces interventions sont des processus physiques susceptibles de modifier l'état du système (approche substantielle). Un système possédant une entropie élevée nécessite pour sa description un long message (approche sémantique) qui ne peut être compressé en raison de la dispersion de ces constituants (par ex. avec un message « toutes les molécules sont à droite »). Cette dispersion demeure la même, qu'un observateur la mesure ou non. Cependant, *si* un quelconque observateur souhaitait suivre le mouvement de toutes les molécules d'un système macroscopique, sans doute qu'après un certain temps la quantité d'intervention et d'information nécessaire à une description fidèle deviendrait ingérable et qu'il voudra plutôt décrire ce système par l'entropie statistique laissant de côté les menus détails microscopiques. Le concept d'entropie statistique n'est pas subjectif pour autant.

La conclusion de la discussion quant à l'interprétation du concept d'entropie peut sembler insatisfaisante, mais elle est en fait riche d'enseignements, car si le travail d'interprétation revient pour une large part à justifier le choix de concepts primitifs afin de mettre en lumière le concept interprété, cette multitude d'interprétations peut donc être considérée comme symptomatique du caractère fondamental du concept d'entropie. Les différentes justifications visant à asseoir la validité d'une interprétation particulière permettent en retour d'établir et de clarifier plusieurs liens qu'entretient ce concept avec d'autres tout aussi fondamentaux, comme ceux de probabilité et d'information. En prêtant attention aux subtilités de ces justifications on en vient à distinguer différentes approches (par ex. les approches substantielle et sémantique) ainsi que différentes positions épistémologiques orientant l'interprétation. On a vu toutefois que l'interprétation de la dispersion était une candidate évitant plusieurs des écueils typiques à ce genre d'entreprise, se rapprochant davantage de la représentation originale de la thermodynamique, et qu'en général les autres interprétations en sont une conséquence plutôt que l'inverse.

La stratégie explicative de l'entropie statistique est double en ce qu'elle présente des différences entre les cas d'équilibre et les cas d'augmentation d'entropie, qui nous ramènent aux fondements mêmes de la mécanique statistique. En tentant de définir une gamme de possibles en accord avec nos mesures empiriques, cette stratégie ne peut faire l'économie du concept de probabilité. Que celui-ci soit indispensable n'implique pas toutefois que sa mise en application soit exempte de problèmes. Car l'attribution empirique d'une valeur de probabilité ne peut recevoir la garantie d'une inférence à partir d'une fréquence relative ni d'un postulat de symétrie se ramenant au principe d'indifférence. Dans le cas de l'équilibre, le raisonnement est plutôt simple et cherche un état microscopique, c'est-à-dire formulé à partir des variables de la mécanique, qui puisse légitimement être un substitut statistique qui serait temporellement invariant, puisqu'un état d'équilibre est défini temporellement invariant. Puisqu'il existe une infinité de distributions compatibles, la sélection de la distribution recherchée consiste en une distribution uniforme. Cette distribution uniforme maximise aussi la valeur d'une entropie préalablement définie. On peut alors justifier ce choix par des arguments de symétrie avec le principe de raison insuffisante ou principe d'indifférence afin de conjuguer avec notre ignorance, ou encore on peut se tourner vers l'ergodicité, qui ne fonctionne que lorsque les temps sont infinis.

Dans le cas de l'augmentation d'entropie, aucune fonction de l'espace des phases ne peut répondre à cette exigence d'accroissement monotone dans le temps. La seule solution qui semble être envisagée, outre le recours à des conditions initiales particulières, est la méthodologie du grenage grossier. Elle permet d'éviter le problème de l'invariance temporelle (on se rappelle ici que l'on souhaite que l'entropie change, augmente) découlant du théorème de Liouville, mais contredit la mécanique. Si cette méthodologie peut sembler arbitraire et subjective, elle reflète justement la « grossièreté » de nos interventions macroscopiques : on ne peut raisonnablement prétendre qu'une mesure de température est sensible à l'influence d'une molécule individuelle. Elle s'accorde de plus avec nos observations, *si* l'entropie combinatoire ou macroscopique est adoptée. Car, pour la vaste majorité des conditions initiales, un système isolé de son environnement évolue de manière déterministe vers un micro-état se trouvant dans une région de l'espace des phases exemplifiant une valeur maximale selon ces concepts d'entropie statistique. On peut alors objecter que ces observations sont impertinentes car, au départ, on cherche à les expliquer. On touche là à un problème de taille. Car on pourrait aussi soutenir que l'augmentation d'entropie tant recherchée aurait été introuvable, parmi l'infinité des possibilités de l'espace des phases, n'eût été de l'observation, de l'entropie thermodynamique. L'idée n'est donc pas tant qu'un macro-état évolue d'un état peu probable à un état plus probable. Mais plutôt qu'il soit très probable, par une définition synchronique, que le macro-état corresponde à un micro-état qui évoluera de manière déterministe vers un micro-état d'équilibre. Il s'agit donc d'une explication réductive partielle. Une intervention extérieure peut faire passer le système d'un état d'équilibre à un état de non-

équilibre, mais presque invariablement le micro-état correspondant à cet état de non-équilibre évoluera vers un état d'équilibre correspondant au plus grand volume dans l'espace des phases.

7.5 L'émergence méthodologique

Le concept d'émergence fait souvent écho à ce qui paraît imprévu, inattendu, nouveau. Ce concept se laisse aussi résumer par le slogan « le tout est supérieur à la somme des parties », mais en quoi ce tout serait « supérieur » et dans quel sens il serait « soudain » est bien entendu matière à réflexion et à débat. Un tout peut être supérieur à la somme de ses parties s'il ne peut être expliqué par un ensemble de concepts, lois et principes propres à ses parties. Mais on n'est guère avancé si le sens de l'explication n'est pas précisé et si cet ensemble de concepts, lois et principes n'est en réalité pas propres aux parties, mais fait plutôt référence à divers niveaux d'organisations.

Le concept d'émergence est un concept compromissaire et relationnel. Il a d'abord été élaboré, dans la littérature philosophique moderne, afin d'obtenir un compromis entre deux ontologies extrêmes mutuellement exclusives : une ontologie moniste et matérialiste, essentiellement réductionniste, où il n'y aurait que des éléments matériels et leurs propriétés, et une ontologie dualiste, donc antiréductionniste, où il y aurait la matière mais aussi l'esprit. Ainsi, bien qu'il n'y ait pas de consensus à son égard, le concept d'émergence est généralement reliée au phénomène qui apparaît à partir de et dépend d'un phénomène plus élémentaire tout en étant dans une certaine mesure autonome de cette même base élémentaire. Mais selon la conception des niveaux d'organisation adoptée, il n'est pas besoin de définir une base élémentaire. Il importe alors de préciser cette relation entre entités (par ex. en termes causalistes) et ces entités mêmes, les relata (propriétés, concepts, processus, etc.) Il s'agit donc de concilier une certaine relation de dépendance, un « enracinement » (« *goundedness* »), avec une relative indépendance, une certaine autonomie. La relation entre ce qui émerge et à partir de quoi cela émerge est souvent explicitée comme une relation entre un tout et ses parties, bref en termes méréologiques.

La conceptualisation de l'émergence s'appuie ici sur deux critères proposés par Huneman (2008) : l'adéquation empirique et la non-trivialité. Les relata retenus sont les propriétés, et ce pour trois raisons. D'abord, il s'agit de relata plus généraux car un processus est un ensemble diachronique de propriétés et un tel ensemble relativement stable peut donner lieu à des lois. Ensuite, ils se prêtent mieux à la conception des niveaux d'organisation adoptée, lesquels sont interprétés comme des maxima locaux de régularité et de prédictibilité de certaines propriétés de composition et d'organisation de la matière. Enfin, la notion de propriété s'applique aussi bien au contexte épistémologique qu'ontologique. En effet, le cadre réaliste adopté autorise les inférences de prémisses épistémologiques à des conclusions ontologiques, car elles peuvent être utiles à la systématisation de nos connaissances, sans que cela signifie que toute inférence du genre soit légitime ou justifiée, ou même désirable.

L'émergence peut ainsi être de type épistémologique ou de type ontologique, mais cette dernière est aux prises avec deux problèmes de taille. L'un de ces problèmes est de distinguer les propriétés émergentes des propriétés ordinaires, comme avec la notion de nouveauté, et pour ce faire il est, semble-t-il, impossible de faire l'économie des théories scientifiques. L'autre problème est de spécifier la relation entre niveaux, car si l'émergence correspond à la structuration effective des processus naturels, elle doit expliciter ce qui relie les propriétés basiques et celles émergentes, autrement sa portée n'est que stipulée. Encore une fois, cette relation, causale ou autre, est généralement celle entre explanans et explanandum. En revanche, l'un de ses avantages est de détacher partiellement la notion d'émergence de celle d'explication en ceci qu'un phénomène émergent ne se limite pas à un phénomène inexpliqué. L'émergence ontologique est alors définie comme l'occurrence d'une propriété qualitativement nouvelle, c'est-à-dire une propriété qui n'est pas possédée par les parties d'un tout. De sorte qu'un phénomène émergent peut être expliqué par une explication faisant référence aux entités reliées aux parties (propriétés, concepts et lois), une explication de type réductif donc, sans que ce phénomène ne s'y réduise totalement. Car, entre autres, malgré une telle explication, différents modes d'accès spécifiques à certaines propriétés peuvent donner lieu à des avantages cognitifs ou représentationnels menant soit à les considérer comme vertus épistémologiques, soit à les interpréter selon le critère de robustesse et donc à leur conférer le statut d'entité réelle *et* nouvelle.

Mais l'émergence ontologique n'est pas nécessairement un objectif à poursuivre à tout prix, la destination de l'émergence épistémologique une fois ses bases bien établies. Celle-ci est à l'œuvre lorsque la relation entre relata ne peut être précisée et que ces relata ne peuvent être considérés robustes. Ces conditions ne sont évidemment pas toujours satisfaites. Il arrive que les desiderata d'une explication soient rencontrés sans que l'explanandum soit considéré « réduit », car bien souvent des hypothèses auxiliaires et des concepts dits « trans-niveaux » ou faisant référence à l'environnement, et qui ne sont pas exclusifs à un certain niveau, sont mobilisés dans cette explication. De plus, le recours à des idéalizations est compatible avec l'explication en général, mais ce n'est pas le cas de l'émergence ontologique. L'exemple de l'entropie de Gibbs, faisant appel à des ensembles infinis de sous-systèmes, est à cet égard probant. Mais puisque le concept d'émergence ne peut se dissoudre en une position moniste, une réduction totale au sens d'identité ou d'élimination doit être écartée. Car il y aurait alors la possibilité d'une réduction de type dérivationnel où les entités au niveau supérieur sont déduites ou dérivées des entités au niveau inférieur.

Cependant, il est désirable, autant que possible, de satisfaire le critère de robustesse de Wimsatt, même si cela ne semble pas toujours possible. C'est pourquoi lier la notion d'émergence à celle d'explication réductive implique de lier cette notion à nos moyens épistémiques et donc à des contingences. Ce qui est considéré « émergent » ne le serait pas nécessairement demain. À la rigueur,

ce n'est pas rédhibitoire car c'est toujours le cas, ou presque, et l'incertitude ne devrait pas nous empêcher d'avancer. Alors, même lorsqu'un phénomène est considéré émergent, la recherche d'explication de type réductive ne devrait pas être interrompue.

C'est pourquoi l'émergence peut aussi être considérée comme un principe *méthodologique* qui exhorte à la recherche d'explication réductionniste autant que possible, sans multiplier les entités théoriques et sans pour autant coucher les phénomènes sur un lit de mystère, tombant par exemple dans un pluralisme exacerbé des niveaux d'organisation selon leurs différents modes d'accès. Affirmer qu'un phénomène est émergent, c'est affirmer qu'il existe un lien modal (par ex. la survenance) entre des entités à des niveaux d'organisation différents sans qu'une explication réductive ne soit possible (par ex. il n'y pas de température sans atome). En ce sens, identifier un cas d'émergence revient à caractériser une situation théorique particulière où différents modes d'accès permettent de définir des niveaux d'organisation sans qu'il n'y ait de réduction totale. Ceci dit, comme il n'y a pas de moyen suffisamment consensuel de juger de la valeur des modes d'accès, même la possibilité d'une explication réductive n'élimine pas totalement la légitimité d'une affirmation niant le succès de la réduction pour des raisons cognitives ou représentationnelles ou pragmatiques. C'est ainsi réaffirmer la confiance en des théories opérant à des niveaux d'organisation différents mais entretenant tout de même des liens entre elles, ce que montre bien le concept d'entropie statistique.

Malgré des théories scientifiques et des modes d'accès fiables, on se trouve dans l'impasse (en pratique ou en principe) d'en offrir une explication jugée satisfaisante (généralement, mais pas nécessairement, respectant le canon réductionniste). C'est-à-dire, malgré certaines explications valables, il reste un programme de recherche réductionniste aussi valables mais qui n'est pas abouti. Par exemple, le phénomène de la vie est largement inexplicable, au sens réductionniste du terme, par la physique, même si l'on ne doute pas de l'existence d'organismes vivants, et même si des progrès importants sont en cours en biochimie et biophysique. L'explication particulière que peut apporter le concept d'émergence vient de la mise en lumière d'une situation théorique où est affirmée l'existence de certaines entités à des niveaux d'organisation différents, malgré l'impossibilité d'expliquer totalement les uns par les autres et malgré la confiance que l'on peut avoir en leur existence et leurs relations mutuelles. Il ne s'agit donc pas d'une explication du phénomène qualifié d'émergent, mais plutôt de l'explication des défis théoriques qu'il pose et de la méthodologie proposée.

7.6 L'entropie comme explanandum et explanans de l'émergence

On a vu que l'entropie est généralement considérée comme une propriété émergente, tandis que l'émergence de certaines structures organisées serait le résultat d'une dissipation d'entropie. Ainsi, l'émergence est parfois présentée comme ce qui expliquerait l'entropie alors que l'entropie expliquerait l'émergence. Tels quels, ces deux énoncés ne peuvent être tous deux vrais. Face à la polysémie déconcertante des concepts d'émergence, d'entropie et d'explication, je soutiens que cet apparent paradoxe peut être résolu formellement ainsi : l'émergence (en un sens *A*) explique (en un sens *B*) l'entropie (en un sens *C*) et l'entropie (en un sens *D*) explique (en un sens *E*) l'émergence (en un sens *F*). Cette solution semble aux premiers abords peu orthodoxe mais elle est fait monnaie courante en philosophie, bien qu'elle ne soit généralement pas explicitée. La solution revient donc à préciser *A*, *B*, *C*, *D*, *E*, et *F*.

Dans le cas de l'entropie comme explanandum de l'émergence (*A*, *B* et *C*), ce qui émerge est l'irréversibilité comme propriété essentielle de l'entropie thermodynamique, mais l'émergence ne peut être synonyme de non-explicabilité. J'ai montré trois possibilités où l'émergence peut expliquer l'entropie thermodynamique. D'abord selon un sens fort avec l'émergence ontologique, où est défini un élément de la structure effective des processus naturels, ou un processus réel, produisant des phénomènes émergents comme l'entropie. Identifier un tel processus n'est pas une mince tâche et pour que la stratégie soit effective, elle ne peut se limiter à stipuler les conditions sous lesquelles un phénomène doit être considéré « émergent ». Ensuite, selon un sens que l'on peut qualifier d'intermédiaire d'après ce que j'ai appelé l'émergence méthodologique. Il y a dans ce cas possibilité d'explication réductive sans réduction totale, et face à cette diversité de niveaux d'organisation et de modes d'accès, l'émergence méthodologique suspend son jugement lorsque plusieurs modes d'accès définissent des niveaux d'organisation différents sans que leur valeur épistémologique ne soit clairement établie. Enfin, selon un sens faible, comme sa désignation comme membre d'une classe d'émergents qui permet donc de définir trivialement une classe de référents pouvant s'y appliquer et une classe de référents ne le pouvant pas. Il y a là une explication au sens faible puisque la gamme des possibles ou l'espace des phases est réduite et une explication plus précise demeure possible et même souhaitée.

Dans le cas de l'entropie comme explanans de l'émergence (*D*, *E*, et *F*), il faut distinguer l'approche substantielle de l'approche analogique. Dans le premier cas, l'entropie renvoie à une propriété macroscopique robuste et autonome pouvant être mobilisée au sein d'un explanans de l'émergence de

nouvelles structures complexes. La stratégie généralisée du principe entropique stipule qu'un système suffisamment complexe en interaction avec son environnement par des échanges d'énergie et d'entropie adopte une configuration particulière qui a comme conséquence un écoulement facilité et une diminution de la génération d'entropie. La configuration que le système développe est une organisation structurale en ce sens où les sous-unités entretiennent des relations mutuelles selon une fonctionnalité globale, qui, selon la théorie constructale, consiste en l'optimisation des contraintes locales en vue de faciliter l'écoulement des flux en présence. En définissant un processus, soit celui de l'optimisation de la facilitation des flux, l'occurrence de organisation structurale peut recevoir une explication causale-nomique, car le processus est encadré globalement par le principe entropique et plusieurs facteurs concourent à cette occurrence à la manière interventionniste. C'est pourquoi la complexité de la configuration n'est pas un objectif et elle n'est pas maximisée, mais plutôt optimisée selon les contraintes locales. De plus, elle n'est pas déterminée *a posteriori* par une règle algorithmique arbitraire et ne dépend pas de l'éventuel effet de surprise d'un quelconque observateur. Cette configuration nouvelle par rapport aux sous-unités (parties) dépend d'elles et de leur organisation sans toutefois s'y réduire et entretient avec elles un lien modal (constitutif). Il y a donc émergence en ce sens. Aussi, ce processus est à la base du maintien de la configuration et éventuellement de son évolution. Le maintien de cette structure peut alors être interprété comme la « survie » du système complexe. Qu'il soit physique ou biologique, celui-ci se voit subsumé sous un principe, la loi constructale, et c'est pourquoi le phénomène du vivant perd en quelque sorte son statut spécial par rapport aux phénomènes physiques.

Dans le cas de l'approche analogique l'entropie exemplifie la réalisabilité multiple et peut être mobilisée, modulo une justification, au sein d'un explanans de l'émergence de propriétés à des niveaux supérieurs. La stratégie explicative est de type déductivo-nomologique en ce sens où l'explanans est un principe entropique analogue. L'hypothèse de base est que l'hypothèse que tout système exemplifiant le concept d'entropie possède certaines propriétés et que si *tel* système exemplifie un analogue du concept d'entropie, alors *donc*, *ce* système possède certaines propriétés, à l'instar de ce qui découle du concept original. Plus souvent qu'autrement, la ressemblance permettant d'inférer cette exemplification du concept est la réalisabilité multiple d'un tout et de ses parties, et cette propriété est celle de la non-décroissance et indirectement celle de l'irréversibilité. Pour que cette stratégie soit effective, il faut d'abord que l'analogie soit justifiée, que les similitudes identifiées dans des contextes différents permettent d'inférer au comportement semblable voire identique du système auquel est appliqué le concept d'entropie analogue. Il faut ensuite, et surtout, que les caractéristiques de l'entropie sur lesquelles on se base pour inférer ce comportement soient bel et bien celles de l'entropie. Mais expliquer l'augmentation d'une entropie analogue à partir de certaines caractéristiques tout aussi

analogues sur la base que l'entropie originale augmente, est nul et non avvenu s'il s'agit d'un concept d'entropie ne pouvant exemplifier une non-décroissance (augmentation), comme le concept statistique d'entropie. Il faudrait alors que l'explication fasse valoir des hypothèses auxiliaires comme par exemple celle du grenage grossier, exploitée afin de montrer qu'il y a bien non-décroissance.

7.7 Multi-, inter- et transdisciplinarité

Les termes multi-, inter- et transdisciplinarité font désormais partie du jargon à la mode, si je puis dire. Mais encore une fois, et l'on devrait s'y attendre au terme de ce périple, leur définition est loin d'être arrêtée. Or, un examen attentif de diverses récupérations conceptuelles (comme on l'a vu avec le concept d'entropie) contribue à clarifier et à préciser ces définitions.

Par la généralisation de la stratégie explicative du principe entropique où l'application souvent métaphorique implique l'intervention de plusieurs disciplines, parfois sans lien évident, on peut affirmer que le concept d'entropie est un cas de multidisciplinarité. Mais déjà on constate qu'une coexistence sans rapport mutuel de théories mono-disciplinaires s'apparente davantage à une idéalisation qu'à la norme.

Cette stratégie explicative permet aussi d'apporter des explications importantes, malgré des défis qui ne sont pas négligeables, dans des disciplines diverses, comme la physiologie (approche substantielle) ou la génétique des populations (approche analogique), exemplifiant plutôt l'interdisciplinarité, comme l'établissement de relations durables entre disciplines.

Enfin, l'application et la validation de cette stratégie généralisée, pour qu'elle soit vraiment effective, exigent une collaboration accrue entre disciplines permettant à terme de les transcender. Les difficultés d'interprétation du concept d'entropie, qui demandent l'intervention de plusieurs disciplines, révèlent pourtant quelque chose de fondamental dans ce concept, car il est dérivé de plusieurs façons et il exprime une impossibilité dont la corroboration ne fait plus de doute. Et les diverses récupérations par analogies formelles, montrent que plusieurs disciplines concourent à établir des bases de stratégies, de méthodes et de concepts communs, illustrant ainsi la transdisciplinarité (voir Alvargonzález 2011).

Le chemin parcouru pour faire sens du concept d'entropie a ainsi autant de valeur que ce qu'on y trouve à la fin. L'inter- et la transdisciplinarité poursuivent les objectifs de la révolution scientifique mais pas d'autres moyens. Déjà la science pascalienne se considérait comme une entreprise collective. Sans se bercer d'illusions, il est en outre permis d'espérer : en encourageant des rapprochements théoriques et disciplinaires, l'unification des connaissances pourrait contribuer, comme on l'a souvent souhaité, et comme le souhaitait Snow dans son appel au dialogue entre les « deux cultures », à l'union des humains.

*To think I did all that;
And may I say - not in a shy way,
“Oh no, oh no not me,
I did it my way”.*

*For what is a man, what has he got?
If not himself, then he has naught.
To say the things he truly feels;
And not the words of one who kneels.
The record shows I took the blows -
And did it my way!*

Yes, it was my way.

Frank Sinatra

8 Bibliographie

1. Adami, C. (2002). "What is Complexity?". *BioEssays*, 24, 1085-1094.
2. Agazzi, E., & Montecucco, L. (2002). Complexity and Emergence. In. Singapour: World Scientific Publishing Co.
3. Ainsworth, P. M. (2012). "The Gibbs Paradox and the Definition of Entropy in Statistical Mechanics". *Philosophy of Science*, 79, 542-560.
4. Albert, D. Z. (2000). *Time and chance*. Cambridge: Harvard University Press.
5. Amiot, P. (2006). Thermodynamique. Note de cours. In *Département de physique, de génie physique et d'optique*. Québec: Université Laval.
6. Amundson, R. (1994). "Two Concepts of Constraint: Adaptationism and the Challenge from Developmental Biology". *Philosophy of Science*, 61, 556-578.
7. Amundson, R., & Lauder, G. V. (1994). "Function Without Purpose: The Uses of Causal Role Function in Evolutionary Biology". *Biology and Philosophy*, 9.
8. Aristote. (2005). *La physique*. Paris Vrin.
9. Asimov, I. (1956, 2014). The last question. In: Internet.
10. Aspect, A. (1984). Au crible de l'expérience. In S. Déligéorge (Ed.), *Le monde quantique*. Paris: Éditions du Seuil.
11. Atkins, P. (2007). *Four laws that drive the Universe*. Oxford: Oxford University Press.
12. Atkins, P. W. (1984, 1987). *Chaleur et désordre. Le deuxième principe de la thermodynamique*. Paris: Pour la science - Belin.
13. Atlan, H. (1972, 2006). *L'organisation biologique et la théorie de l'information*. Paris: Seuil.
14. Augros, R., & Stanciu, G. (1987, 1995). *Le monde vivant, une sagesse à découvrir ; discussion de conceptions de l'être humain et du monde vivant*. Sainte-Foy: I.A.R.
15. Baas, N. A., & Emmeche, C. (1997). "On emergence and explanation". *Intellectica*, 2, 67-83.
16. Bach, A. (1990). "Boltzmann's Probability Distribution of 1877". *Archives for the History of Exact Sciences*, 41, 1-40.
17. Balian, R. (2003). Entropy : a protean concept. In *Séminaire Poincaré 2* (pp. 13-27).
18. _____. (2005). "Information in statistical physics". *Studies in History and Philosophy of Modern Physics*, 36, 323-353.
19. _____. (2013). "La longue élaboration du concept d'énergie". *Académie des sciences : Histoire des sciences*, 2013.
20. Balibar, F. (1999). Énergie. In D. Lecourt (Ed.), *Dictionnaire d'histoire et de philosophie des sciences*. Paris: Presses universitaires de France.
21. Ball, P. (1999). *The Self-Made Tapestry : Pattern formation in nature*. Oxford: Oxford University Press.
22. Barberousse, A. (2002). *La mécanique statistique : de Clausius à Gibbs*. Paris: Belin.
23. Barreau, H. A. R.-E. (2009). Du temps physique au temps cosmologique: le rétablissement de la flèche du temps. In *HAL - Sciences de l'homme et de la société*. Cerisy-la-Salle: Internet.
24. Barrett, M., & Sober, E. (1992). "Is Entropy Relevant to the Asymmetry Between Retrodiction and Prediction?". *British Journal for the Philosophy of Science*, 43, 141-160.
25. Barton, N. H., & Coe, J. B. (2009). "On the application of statistical physics to evolutionary biology". *Journal of Theoretical Biology*, 259, 317-324.
26. Batterman, R. (2009). Emergence in physics. In E. Craig (Ed.), *Routledge Encyclopedia of Philosophy*. Londres: Routledge.
27. _____. (2002). *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence* (Routledge Encyclopedia of Philosophy; Oxford: Oxford University Press).
28. _____. (2000). "Multiple Realizability and Universality". *British Journal for the Philosophy of Science*, 51, 115-145.
29. Battino, R., Strong, L. E., & Wood, S. E. (1997). "A Brief History of Thermodynamics Notation". *Journal of Chemical Education*, 74, 304-305.
30. Baubérot, J. (1990). *La laïcité, quel héritage? De 1789 à nos jours*. Genève: Labor et Fides.

31. Bealer, G. (1996). "'A Priori' Knowledge and the Scope of Philosophy". *Philosophical Studies*, 81, 121-142.
32. Bedau, M. A. (2008). Downward Causation and Autonomy in Weak Emergence. In M. A. Bedau & P. Humphreys (Eds.), *Emergence: Contemporary readings in philosophy of science*. Cambridge: MIT Press.
33. Bedau, M. A., & Humphreys, P. (2008). Emergence: Contemporary Readings in Philosophy and Science. In. Cambridge: The MIT Press.
34. Bejan, A. (1998). "Constructal theory: from thermodynamics and geometric optimization to predicting shape in nature". *Energy Conservation Management*, 39, 1705-1718.
35. _____. (2006). *Advanced engineering thermodynamics*. Hoboken: John Wiley & Sons.
36. Bejan, A., & Zane, J. P. (2012). *Design in Nature: how the constructal law governs evolution in biology, physics, technology and social organization*. New York: Doubleday.
37. Bell, J. S. (1987). *Speakable and Unsayable in Quantum Mechanics*. Cambridge: Cambridge University Press.
38. Bellac, M. L., Mortessagne, F., & Batrouni, G. G. (2004). *Equilibrium and non-equilibrium statistical thermodynamics*. Cambridge: Cambridge University Press.
39. Belnap, N. (2009). Notes on the Science of Logic. In. Pittsburgh: University of Pittsburgh.
40. Belot, G., & Earman, J. (1999). From metaphysics to physics. In J. Butterfield & C. Pagonis (Eds.), *From physics to philosophy* (pp. 920-959). Cambridge: Cambridge University Press.
41. Ben-Menahem, Y. (2000). Idealization. In W. H. Newton-Smith (Ed.), *A companion to the philosophy of science*. Oxford: Blackwell.
42. Ben-Naim, A. (2008). *Statistical Thermodynamics Based on Information : A Farewell to Entropy*. Singapore: World Scientific Publishing.
43. Benkirane, R. (2002, 2013). La complexité, vertiges et promesses. In. Paris: Le pommier.
44. Bennett, C. H. (2003). How to Define Complexity in Physics, and Why. In N. H. Gregersen (Ed.), *From Complexity to Life: On the Emergence of Life and Meaning*. Oxford: Oxford University Press.
45. Berenda, C. W. (1953). "On Emergence and Prediction". *The Journal of Philosophy*, 50, 269-274.
46. Beretta, G. P. (1986). Intrinsic Entropy and Intrinsic Irreversibility for a Single Constituent of Matter: Broader Kinematics and Generalized Nonlinear Dynamics. In G. T. Moore & M. O. Scully (Eds.), *Frontiers of Nonequilibrium Statistical Physics*: Plenum.
47. Beretta, M. (2001). Institutionnalisation et professionnalisation. In M. Blais & E. Nicolaïdis (Eds.), *L'Europe des sciences - constitution d'un espace scientifique*. Paris: Éditions du Seuil.
48. Berry, S. (1995). "Entropy, Irreversibility and Evolution". *Journal of Theoretical Biology*, 175, 197-202.
49. Bickhard, M. H. (2011). Metaphysical and Epistemological Issues in Complex Systems. In C. Hooker (Ed.), *Handbook of the Philosophy of Science* (Vol. 10, pp. 105-136). Oxford: Elsevier.
50. Biedenbarn, L. C., & Solem, J. C. (1995). "A Quantum-Mechanical Treatment of Szilard's Engine: Implications for the Entropy of Information". *Foundations of Physics*, 25, 1221-1229.
51. Birkhoff, G. D. (1931). "Proof of the ergodic theorem". *Proceedings of the National Academy of Sciences*, 18, 279-282.
52. Bishop, R. C. (2011). Metaphysical and Epistemological Issues in Complex Systems. In C. Hooker (Ed.), *Handbook of the Philosophy of Science* (Vol. 10, pp. 105-136). Oxford: Elsevier.
53. Bitbol, M. (1996, 2003). *Mécanique quantique : une introduction philosophique*. Paris: Champs-Flammarion.
54. _____. (2007). "Ontology, matter and emergence". *Phenomenology and cognitive sciences*, 6, 293-307.
55. _____. (2012). "Downward causation without foundations". *Synthese*, 185, 233-255.
56. Blondel, C. (1998). "Les physiciens français et l'électricité industrielle à la fin du XIXe siècle". *Physis, Rivista internazionale di storia della scienza*, 35, 245-271.
57. Boksenbojm, E., Wynants, B., & Jarzynski, C. (2010). "Nonequilibrium thermodynamics at the microscale: Work relations and the second law". *Physica A*, 389, 4406-4417.
58. Boltzmann, L. (1872). Further Studies on the Thermal Equilibrium of Gas Molecules. In S. G. Brush (Ed.), *The Kinetic Theory of Gases*. London: Imperial College Press.
59. Boogerd, F. C., Bruggeman, F. J., Richardson, R. C., Stephan, A., & Westerhoff, H. (2005). "Emergence and its Place in Nature: A Case Study of Biochemical Networks". *Synthese*, 145, 131-164.
60. Bouchard, F. (2008). "Causal processes, fitness, and the differential persistence of lineages". *Philosophy of Science*, 75, 560-570.
61. _____. (2011). "Darwinism without populations: a more inclusive understanding of the 'Survival of the Fittest'". *Studies in History and Philosophy of Biological and Biomedical Sciences*, 42, 106-114.

62. Bouchard, F., & Rosenberg, A. (2004). "Fitness, Probability and the Principles of Natural Selection". *British Journal for the Philosophy of Science*, *axh406*, 693-712.
63. Boyer, C. B. (1968, 1991). *A History of Mathematics*. New York: John Wiley & sons.
64. Brading, K., & Castellani, E. (2003). *Symmetry in Physics - Philosophical Reflexions*. Cambridge: Cambridge University Press.
65. Brandon, R. N. (1997). "Does Biology Have Laws? The Experimental Evidence". *Philosophy of Science (Proceedings)*, 544-457.
66. Brenner, A. (2011). *Raison scientifique et valeurs humaines : essai sur les critères du choix objectif*. Paris: Presses Universitaires de France.
67. Bricmont, J. (2001). Bayes, Boltzmann and Bohm : Probabilities in Physics. In J. Bricmont, D. Dürr, M. C. Galavotti, G. Ghirardi, F. Petruccione & N. Zanghi (Eds.), *Chance in Physics. Foundations and Perspectives*. Berlin: Springer.
68. Bricmont, J., Dürr, D., Galavotti, M. C., Ghirardi, G., Petruccione, F., & Zanghi, N. (2001). Chance in Physics : Foundations and Perspectives. In *Lecture Notes in Physics, Vol. 574*. Berlin: Springer.
69. Bridgman, P. W. (1941, 1943). *The Nature of Thermodynamics*. London: Oxford University Press.
70. Brigandt, I., & Love, A. (2008). Reductionism in biology. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*: Internet.
71. Brin, M., & Stuck, G. (2003). Introduction to Dynamical Systems. In Internet (Ed.). Cambridge: Cambridge University Press.
72. Brissaud, J.-B. (2005). "The meaning of Entropy". *Entropy*, 7, 68-96.
73. Brooks, D. R., & Wiley, E. O. (1988). *Evolution as Entropy: Toward a Unified Theory of Biology*. Chicago: University of Chicago Press.
74. Brown, H. R., Myrvold, W., & Uffink, J. (2009). "Boltzmann's H-theorem, its discontents, and the birth of statistical mechanics". *Studies in History and Philosophy of Modern Physics*, 40, 174-191.
75. Brown, H. R., & Uffink, J. (2001). "The Origins of Time-Asymmetry in Thermodynamics: The Minus First Law". *Studies in History and Philosophy of Modern Physics*, 32, 525-538.
76. Brunhes, B. (1909, 1991). *La dégradation de l'énergie*. Paris: Champs - Flammarion.
77. Brunold, C. (1930). *L'entropie: son rôle dans le développement historique de la thermodynamique*. Paris: Masson et cie.
78. Brush, S. G. (1976). *The Kind of Motion we Call Heat. A History of the Kinetic Theory of Gases in the 19th Century* (Vol. I : Physics and the Atomists). New York: North-Holland Publishing.
79. _____. (2003). *The Kinetic Theory of Gases*. London: Imperial College Press.
80. Bub, J. (2007). "Quantum probabilities as degree of belief". *Studies in History and Philosophy of Modern Physics*, 38, 232-254.
81. Bunge, M. (1959, 1963). *Causality. The place of the causal principle in modern science*. Cleveland: Meridian Books.
82. _____. (1966). "Mach's critique of newtonian mechanics". *American Journal of Physics*, 34, 585-596.
83. _____. (1967). *Foundations of physics*. New York: Springer-Verlag.
84. _____. (1967). "Analogy in quantum theory : from insight to nonsense". *British Journal for the Philosophy of Science*, 18, 265-286.
85. _____. (1967, 1998). *Philosophy of science* (Vol. 1 et 2). New Jersey: Transaction.
86. _____. (1968). "Physical Time: The Objective and Relational Theory". *Philosophy of Science*, 35, 355-388.
87. _____. (1997). "Mechanism and Explanation". *Philosophy of the Social Sciences*, 27, 410-465.
88. _____. (2000). "Energy: Between Physics and Metaphysics". *Science and Education*, 9, 457-461.
89. _____. (2003). *Emergence and convergence*. Toronto: Toronto University Press.
90. _____. (2006). *Chasing reality - strife over realism*. Toronto: Toronto University Press.
91. Burian, R. M. (2005). *The Epistemology of Development, Evolution, and Genetics. Selected Essays*. Cambridge: Cambridge University Press.
92. Butterfield, J. (2011). "Emergence, Reduction and Supervenience: a Varied Landscape". *Foundations of Physics*, 41, 920-959.
93. Butterfield, J., & Pagonis, C. (1999). From physics to philosophy. In (pp. 920-959). Cambridge: Cambridge University Press.
94. Callender, C. (1999). "Reducing Thermodynamics to Statistical Mechanics: The Case of Entropy". *The Journal of Philosophy*, 96, 348-373.

95. _____. (2001). "Taking Thermodynamics Too Seriously". *Studies in History and Philosophy of Modern Physic*, 32, 539-553.
96. _____. (2002). Time, Reality & Experience. In. Cambridge: Cambridge University Press.
97. _____. (2004). "A Collision between Dynamics and Thermodynamics". *Entropy*, 6, 11-20.
98. _____. (2006). Thermodynamic Asymmetry in Time. In S. University (Ed.), *Stanford Encyclopedia of Philosophy*: Internet.
99. Cameron, R. (1985). "A New View of European Industrialization". *The Economic History Review*, 38, 1-24.
100. Campbell, M. K., & Farrell, S. O. (2006). *Biochemistry*. Melmont: Thomson Brooks/Cole.
101. Campisi, M. (2008). "Statistical mechanical proof of the second law of thermodynamics based on volume entropy". *Studies in History and Philosophy of Modern Physics*, 39.
102. Campisi, M. (2008). "On the mechanical foundations of thermodynamics: The generalized Helmholtz theorem". *Studies in History and Philosophy of Modern Physics*, 36.
103. Capek, V., & Sheehan, D. P. (2005). *Challenges to the Second Law of Thermodynamics. Theory and Experiment*. Dordrecht: Springer.
104. Cardwell, D. S. L. (1967). "Some Factors in the Early Development of the Concepts of Power, Work and Energy". *The British Journal for the History of Science*, 3, 209-224.
105. Carnap, R. (1966). *Philosophical foundations of physics*. New York: Basic Books.
106. _____. (1977). *Two Essays on Entropy*. Berkeley & Loas Angeles: University of California Press.
107. Carnot, S. (1824, 1878). *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance*. Paris Gauthier-Villard [Bachelier].
108. Carnot, S. (1878). *Extrait de notes inédites de Sadi Carnot sur les mathématiques, la physique et autres sujets*. Paris: Bachelier.
109. Cartwright, N. (1982). "When explanation leads to inference". *The Journal of Philosophy*, 13, 111-121.
110. Cartwright, N. (1983, 2002). *How the laws of physics lie*. Oxford: Clarendon Press.
111. Cassirer, E. (1910, 1977). *Substance et fonction. Éléments pour une théorie du concept*. Paris: Les Éditions de Minuit.
112. Castoriadis, C. (1999). Technique. In *Dictionnaire de la philosophie*. Paris: Albin Michel. Encyclopaedia Universalis.
113. Caves, C. M., Fuchs, C. A., & Schack, R. (2007). "Subjective probability and quantum certainty ". *Studies in History and Philosophy of Modern Physic*, 38, 255-274.
114. Çengel, Y. A., & Boles, M. A. (2006). *Thermodynamics: An Engineering Approach* (5e ed.): McGraw-Hill.
115. Cercignani, C. (1998). *Ludwig Boltzmann: The Man Who Trusted Atoms*. Oxford: Oxford University Press.
116. Chakravartty, A. (2007). *A Metaphysic for Scientific Realism: Knowing the Unobservable*. Cambridge: Cambridge University Press.
117. Chakravartty, A. (2007). "Structuralism as a form of scienti@c realism". *International studies in the philosophy of science*, 18, 151-171.
118. Chalmers, A. F. (1976, 1982). *Qu'est-ce que la science ?* Paris: La découverte.
119. Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory* Oxford: Oxford University Press.
120. Chang, H. (2004). *Inventing Temperature : Measurement and Scientific Progress*. Oxford: Oxford University Press.
121. Charle, C. (2004). Chapitre 2 : Patterns. In W. Rüegg (Ed.), *University in Europe* (Vol. 3). Cambridge: Cambridge University Press.
122. Clagett, M. (1959, 1962). Critical Problems in the History of Science. In. Madison: University of Wisconsin Press.
123. Clark, P. J. (2001). Statistical Mechanics and the Propensity Interpretation of Probability. In J. Bricmont, D. Dürr, M. C. Galavotti, G. Ghirardi, F. Petruccione & N. Zanghi (Eds.), *Chance in Physics. Foundations and Perspectives*. Berlin: Springer.
124. Clausius, R. (1850). On the motive power of heat, and on the laws which can be deduced from it for the theory of heat In W. F. Magie (Ed.), *The second law of thermodynamics: Memoirs by Carnot, Clausius and Thomson*. New York: Haper & Brothers Publishers.
125. Clausius, R. (1851). On the Moving Force of Heat, and the Laws regarding the Nature of Heat itself which are deducible therefrom. In D. Brewster, R. Taylor, R. Kane & W. Francis (Eds.), *The London,*

- Edinburgh, and Dublin Philosophical Magazine and Journal of Science* (Vol. II). London: University of London.
126. Clausius, R. (1862). On the Application of the Theorem of the Equivalence of Transformations to the Internal Work of a Mass of Mater. In D. Brewster, R. Taylor, R. Kane & W. Francis (Eds.), *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* (Vol. II). London: University of London.
 127. Clausius, R. (1868). *Théorie mécanique de la chaleur*. Paris: Eugène Lacroix.
 128. Clayton, P., & Davies, P. (2006). The reemergence of emergence : the emergentist hypothesis from science to religion. In. Oxford: Oxford University Press.
 129. Coe, L. (1969). "The Nature of Time". *American Journal of Physics*, 37, 810-815.
 130. Coffman, J. A. (2011). Metaphysical and Epistemological Issues in Complex Systems. In C. Hooker (Ed.), *Handbook of the Philosophy of Science* (Vol. 10, pp. 105-136). Oxford: Elsevier.
 131. Cohen, B. (1969). *Les origines de la physique moderne*. Paris: Petite bibliothèque Payot.
 132. Cohen, E. G. D. (2008). "Entropy, Probability and Dynamics". *arXiv:0807.1268v2*, 10 jul.
 133. Cohen-Tannoudji, C., Diu, B., & Laloë, F. (1973, 1998). *Mécanique quantique. Tome 1 et 2*. Paris: Hermann.
 134. Comte, A. (1844, 1974). *Discours sur l'esprit positif*. Paris: Vrin.
 135. Corradini, A., & O'Connor, T. (2010). Emergence in Science and Philosophy. In. New York: Routledge.
 136. Coveney, P. V. (1988). "The Second Law of Thermodynamics: Entropy, Irreversibility and Dynamics". *Nature*, 333, 409-415.
 137. Cropper, W. H. (2001). *Great Physicists. The Life and Times of Leading Physicists from Galileo to Hawking*. New York: Oxford University Press.
 138. Daintith, J. (2009). Biographical Encyclopedia of Scientists. Third Edition. In. Boca Raton: CRC Press.
 139. Darrigol, O. (2003). The Origins of the Entropy Concept. In *Séminaire Poincaré 2* (pp. 1-12).
 140. Darwin, C. (1859, 1993). *The Origin of Species*. New York: Modern Library.
 141. Daumas, M. (1957). *Histoire de la science*. Paris: La Pléiade.
 142. Davey, K. (2008). "The Justification of Probability Measures in Statistical Mechanics". *Philosophy of Science*, 75, 28-44.
 143. Davies, P. (2003). Complexity and the arrow of time. In N. H. Gregersen (Ed.), *From Complexity to Life: On the Emergence of Life and Meaning*. Oxford: Oxford University Press.
 144. Dawkins, R. (1976, 2006). *The selfish gene*. Oxford: Oxford University Press.
 145. de Chardin, T. (1956). *La place de l'homme dans la nature*. Paris: Albin Michel.
 146. de la Cotardière, P. (2004, 2012). Histoire des sciences. In. Paris: Tallandier.
 147. Deacon, T. W. (2003). Three levels of emergent phenomena. In N. C. Murphy & W. R. Stoeger (Eds.), *Evolution and emergence*. Oxford: Oxford University Press
 148. Demetrius, L. (1974). "Demographic Parameters and Natural Selection". *Proceedings of the National Academy of Sciences, USA*, 71, 4645-4647.
 149. Demetrius, L. (1977). "Measures of fitness and demographic stability". *Proceedings of the National Academy of Sciences, USA*, 74, 384-386.
 150. Demetrius, L. (2000). "Thermodynamics and Evolution". *Journal of Theoretical Biology*, 206, 1-16.
 151. Demetrius, L., Legendre, S. p., & Harremöes, P. (2009). "Evolutionary Entropy: A Predictor of Body Size, Metabolic Rate and Maximal Life Span". *Bulletin of Mathematical Biology*.
 152. Denbigh, K. G. (1989). "Note on Entropy, Disorder and Disorganization". *The British Journal for the Philosophy of Science*, 40, 323-332.
 153. Denbigh, K. G. (1989). "The Many Faces of Irreversibility". *The British Journal for the Philosophy of Science*, 40, 501-518.
 154. Depew, D. J. (1986). "Nonequilibrium Thermodynamics and Evolution: a Philosophical Perspective". *Philosophica*, 37, 27-58.
 155. DeRose, K. (1995). "Solving the Skeptical Problem". *The Philosophical Review*, 104, 1-52.
 156. Desjardins, E. (2011). "Reflections on Path Dependence and Irreversibility: Lessons from Evolutionary Biology". *Philosophy of Science*, 78, 724-738.
 157. Deutsch, G. (2008). *The Entropy Crisis*. Singapore: World Scientific Publishing.
 158. Dewar, R. C. (2010). Maximum Entropy Production and Non-equilibrium Statistical Mechanics. In: Internet: <http://homepage.mac.com/bradmarston/Papers/chpt04.pdf>.

159. Dias, P. M. C., Pinto, S. P., & Cassiano, D. H. (1994). "The conceptual import of Carnot's theorem to the discovery of the entropy". *Archives for the History of Exact Sciences*, 49, 135-161.
160. Dieks, D. (2006). The Ontology of Spacetime. In D. Dieks & M. Rédei (Eds.), *Philosophy and Foundations of Physics*. Amsterdam
161. Oxford: Elsevier.
162. Diu, B., Guthmann, C., Lederer, D., & Roulet, B. (1989). *Physique statistique*. Paris: Hermann.
163. Dobbs, B. J. T. (1988). Alchemy and Newton's principle of gravitation. In P. B. Scheurer & G. Debrck (Eds.), *Newton's scientific and philosophical legacy*. Dordrecht: Kluwer Academic Publishers.
164. Dobzhansky, T. (1973). "Nothing in Biology Makes Sense Except in the Light of Evolution". *The American Biology Teacher*, 35, 125-129.
165. Dorato, M. (2005). Why are (most) laws of nature mathematical ? In J. Faye, P. Needham, U. Scheffler & M. Urchs (Eds.), *Nature's Principles*. Amsterdam: Springer.
166. Drossel, B. (2001). "Biological evolution and statistical physics". *Advances in Physics*, 50, 209-295.
167. Dubucs, J. (2002). "Feasibility in logic". *Synthese*, 132, 213-237.
168. Dugdale, J. S. (1996). *Entropy and its Physical Meaning*. London: Taylor & Francis.
169. Dummett, M. (1978). *Truth and Others Enigmas*. Cambridge: Harvard University Press.
170. Duncan, T. L., & Semura, J. S. (2007). "Information Loss as a Foundational Principle for the Second Law of Thermodynamics". *Foundations of Physics*, 37, 1767-1773.
171. Earman, J. (1970). "Who's afraid of absolute space?". *Australasian Journal of Philosophy*, 48, 287-319.
172. Earman, J. (1986). "The Problem of Irreversibility". *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science*, 2: Symposia and Invited Papers, 226-233.
173. Earman, J. (1992). Inference, explanation, and other frustrations : essays in the philosophy of science. In (pp. 226-233). Berkeley: University of California Press.
174. Earman, J. (2010). Laws, Symmetry, and Symmetry Breaking; Invariance, Conservation Principles, and Objectivity. In: Internet.
175. Earman, J., & Rédei, M. (1996). "Why Ergodic Theory Does Not Explain the Success of Equilibrium Statistical Mechanics". *The British Journal for the Philosophy of Science*, 47, 63-78.
176. Eckhardt, W. (2006). "Causal time asymmetry". *Studies in History and Philosophy of Modern Physics*, 37, 439-466.
177. Eddington, A. S. (1928, 1929). *The Nature of the Physical World*. New York: Macmillan Company.
178. Edmonds, B. (1999). *Syntactic Measures of Complexity*. University of Manchester Manchester
179. Edwards, A. W. F. (2000). "The Genetical Theory of Natural Selection". *Genetics*, 154, 1419-1426.
180. Ehrenfest, P., & Ehrenfest, T. (1912, 1990). *The conceptual foundations of the statistical approach in mechanics*. New York: Dover Publications, inc.
181. Eigen, M. (1971). "Selforganization of Matter and the Evolution of Biological Macromolecules". *Die Naturwissenschaften*, 58, 465-523.
182. Einstein, A. (1951). Autobiographical notes. In P. A. Schlipp (Ed.), *Albert Einstein: Philosopher-Scientist*. New York: MJF Books.
183. Einstein, A. (1956). *La relativité*. Paris: Petite bibliothèque Payot.
184. Einstein, A., & Infeld, L. (1953). *L'évolution des idées en physique*. Paris: Petite bibliothèque Payot.
185. Einstein, A., & Infled, L. (1963). *L'évolution des idées en physique*. Paris: Petite bibliothèque Payot.
186. Eisenstaedt, J. (1998). La relativité générale : une révolution ? In *Cahiers d'épistémologie* (Vol. 250). Montréal: Université du Québec à Montréal.
187. Elgin, C. Z. (2004). "True enough". *Philosophical Issues*, 14, 113-131.
188. Emch, G. G. (2005). "Probabilistic issues in statistical mechanics". *Studies in History and Philosophy of Modern Physics*, 36, 303-322.
189. Emmeche, C. (1997). Defining life, explaining emergence. In www.nbi.dk/~emmeche/cePubl/97e.defLife.v3f.html.
190. Ereshefsky, M. (2012). "Homology Thinking". *Biology and Philosophy*, 27, 381-400.
191. Erlichson, H. (1999). "Sadi Carnot, 'Founder of the Second Law of Thermodynamics'". *European Journal of Physics*, 20, 183-192.
192. Evans, D. J., & Searles, D. J. (2002). "The Fluctuation Theorem". *Advances in Physics*, 51.
193. Falk, R. (2009). *Genetic Analysis: A History of Genetic Thinking*. Cambridge: Cambridge University Press.
194. Farmer, J. D., & Geanakoplos, J. (2008). "The Virtues and Vices of Equilibrium and the Future of Financial Economics". *Complexity*, 14.

195. Fast, J. D. (1960, 1962). *Entropy: The Significance of the Concept of Entropy and its Applications in Science and Technology*. New York: McGraw-Hill Books.
196. Feingold, M. (2008). History of Universities. In (Vol. 23). Oxford: Oxford University Press.
197. Feltz, B. (2006). Self-Organization, Selection and Emergence in the Theories of Evolution. In M. C. Bernard Feltz, Philippe Goujon (Ed.), *Self-Organization and Emergence in Life Sciences*. Dordrecht: Springer.
198. Fermi, E. (1936). *Thermodynamics*. New York: Dover Publications.
199. Feuerbach, L. (1841, 1968). *L'essence du christianisme*. Paris: Gallimard.
200. Feynman, R. (1965, 1994). *The character of physical law*. New York: Modern Library.
201. Fink, J. K. (2009). *Physical Chemistry in Depth*. Heidelberg: Springer-Verlag.
202. Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.
203. Fisher, R. A. (1941). "Average excess and average effect of a gene substitution". *Annals of Eugenics*, 11, 53-63.
204. Fisher, R. A. (1990 [1922]). "Average excess and average effect of a gene substitution". *Bulletin of Mathematical Biology*, 52, 297-318.
205. Fleissner, P., & Hofkirchner, W. (1994). Entropy and Its Implications for Sustainability. In J. C. Dragan, M. C. Demetrescu & E. K. Seifert (Eds.), *Implications and Applications of Bioeconomics*. Palma de Mallorca.
206. Fletcher, P. (2002). "A Constructivist Perspective on Physics". *Philosophia Mathematica*, 10, 26-42.
207. Floridi, L. (2004). "On the logical unsolvability of the Gettier problem". *Synthese*, 142, 61-79.
208. Fodor, J. (1968). *Psychological Explanation*. Cambridge: MIT Press.
209. Forge, J. (1980). "The structure of physical explanation". *Philosophy of Science*, 47, 203-226.
210. Fourier, J. (1822). *Théorie analytique de la chaleur*. Paris: Firmin Didot, Père et fils.
211. Fraïssé, F. (2012). *La complexité et les phénomènes : nouvelles ouvertures entre science et philosophie*. Paris: Hermann.
212. Frank, S. A., & Slatkin, M. (1992). "Fisher's Fundamental Theorem of Natural Selection". *TREE*, 7, 92-95.
213. Fridberg, A. (1978). *Sadi Carnot, physicien et les "Carnot" dans l'histoire*. Paris: La pensée universelle.
214. Frigg, R. (2008). "A Field Guide to Recent Work on the Foundations of Statistical Mechanics". *arXiv:0804.0399v1 [cond-mat.stat-mech]*.
215. Frigg, R. (2008). "Chance in Boltzmannian Statistical Mechanics". *Philosophy of Science*, 75, 670-681.
216. Frigg, R., & Werndl, C. (2011). Entropy - A Guide for the Perplexed. In C. Beisbart & S. Hartmann (Eds.), *Probability in Physics*. Oxford: Oxford University Press.
217. Gabbey, W. A., & Herivel, J. W. (1966). "Un manuscrit inédit de Sadi Carnot". *Revue d'histoire des sciences et de leurs applications*, 19, 151-166.
218. Gagnon, M., & Hébert, D. (2000). *En quête de science. Introduction à l'épistémologie*. Montréal: Fides.
219. Gal-Or, B. (1972). "The crisis about the origin of irreversibility and the anisotropy of time". *Science*, 176.
220. Galison, P. (2003). *Einstein's clocks, Poincaré's maps*. New York: W. W. Norton & Company.
221. Gallavotti, G. (1994). "Ergodicity, ensembles, irreversibility in Boltzmann and beyond". *arXiv:chaodyn/9403004v1*.
222. Gallavotti, G. (1999). *Statistical Mechanics. Short treatise*. Rome: Internet.
223. Gallavotti, G., Reiter, W. L., & Yngvason, J. (2008). Boltzmann's Legacy. In. Freiburg: European Mathematical Society.
224. Gauthier, Y. (2009). "The Construction of Chaos Theory". *Foundations of Science*, 14, 153-165.
225. Gaxotte, P. (1975, 1988). *La Révolution française*. Paris: Éditions Complexe.
226. Gayon, J. (1999). Sélection. In D. Lecourt (Ed.), *Dictionnaire d'histoire et philosophie des sciences*. Paris: Presses universitaires de France.
227. Gettier, E. L. (1963). "Is Justified True Belief Knowledge?". *Analysis*, 23, 121-123.
228. Giampietro, M., & Mayumi, K. (2004). Complex Systems and Energy. In C. J. Cleveland (Ed.), *Encyclopedia of Energy* (pp. 617-631): Elsevier.
229. Gibbs, W. (1902, 1960). *Elementary principles in statistical mechanics*. New York: Dovers Publications, inc.
230. Giles, R. (1964). *Mathematical Foundations of Thermodynamics*. Frankfurt: Pergamon Press.
231. Gillett, C. (2003). "The Metaphysics of Realization, Multiple Realizability, and the Special Sciences". *Journal of Philosophy*, 100, 591-603.
232. Gingras, Y., Keating, P., & Limoges, C. (1998). *Du scribe au savant*. Montréal: Boréal.

233. Gleick, J. (2003). *Isaac Newton*. New York: Vintage Books.
234. Glymour, B. (2007). In defense of explanatory deductivism. In J. K. Campbell, M. O'Rourke & H. Silverstein (Eds.), *Causation and explanation*. Cambridge: The MIT Press.
235. Glymour, C. (1970). "On Some Patterns of Reduction". *Philosophy of Science*, 37, 340-353.
236. Gnaiger, E. (1993). "Nonequilibrium thermodynamics of energy transformations". *Pure and Applied Chemistry*, 65, 1982-2002.
237. Goldman, A. I. (1976). "Discrimination and Perceptual Knowledge". *Journal of Philosophy*, 73, 771-791.
238. Goldman, A. I. (1988). "Strong and Weak Justification". *Philosophical Perspectives*, 2, 51-69.
239. Goldstein, S. (2001). Boltzmann's Approach to Statistical Mechanics. In J. Bricmont, D. Dürr, M. C. Galavotti, G. Ghirardi, F. Petruccione & N. Zanghi (Eds.), *Chance in Physics. Foundations and Perspectives*. Berlin: Springer.
240. Goodin, R. E., Pettit, P., & Pogge, T. (1993, 2007). A Companion to Contemporary Political Philosophy. 2nd Edition. In. Malden: Blackwell Publishing.
241. Gould, H., & Tobochnik, J. (2010). *Thermal and Statistical Physics*. Princeton: Princeton University Press.
242. Gould, H. H., & Tobochnik, J. (2010). *Thermal and Statistical Physics*. Princeton: Princeton University Press.
243. Gould, S. J., & Lewontin, R. (1979). "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme". *Proceedings of the Royal Society of London*, 205, 581-598.
244. Grant, E. (1962). "Late Medieval Thought, Copernicus, and the Scientific Revolution". *Journal of the History of Ideas*, 23, 197-220.
245. Grattan-Guinness, I. (1984). "Work for the workers: Advances in engineering mechanics and instruction in France, 1800-1830". *Annals of Science*, 41, 1-33.
246. Greco, J. (1993). "Virtues and vices of virtue epistemology". *Canadian Journal of Philosophy*, 23, 413-432.
247. Greenberger, D., Hentschel, K., & Weinert, F. (2009). Compendium of Quantum Physics : Concepts, Experiments, History and Philosophy. In. Berlin Heidelberg: Springer-Verlag.
248. Gregersen, N. H. (2003). From Complexity to Life: On the Emergence of Life and Meaning. In. Oxford: Oxford University Press.
249. Greiner, W., Neise, L., & Stöcker, H. (1987, 1997). *Thermodynamics and Statistical Physics*. New York: Springer-Verlag.
250. Greven, A., Keller, G., & Warnecke, G. (2010). Entropy. In. Princeton: Princeton University Press.
251. Griffiths, A. J. F., Miller, J. H., Susuki, D. T., Lewontin, R. C., & Gelbart, W. M. (1976, 1996). *An Introduction to Genetic Analysis*, 6th edition. New York: Freeman.
252. Griffiths, P. (2008). Philosophy of Biology. In S. University (Ed.), *Stanford Encyclopaedia of Philosophy*: Internet.
253. Grimm, S. R. (2008). "Explanatory Inquiry and the Need for Explanation". *British Journal for the Philosophy of Science*, 59, 481-497.
254. Guay, A., & Hepburn, B. (2009). "Symmetry and Its Formalisms: Mathematical Aspects". *Philosophy of Science*, 76.
255. Guedj, M. (2006). "Du concept de travail vers celui d'énergie : L'apport de Thomson". *Revue d'histoire des sciences*, 59, 29-50.
256. Guedj, M. (2007). "À propos de l'ouvrage d'Helmholtz Über die Erhaltung der Kraft sur un principe limité de la conservation de l'énergie". *Philosophia Scientiae*, 11, 1-25.
257. Guggenheim, E. A. (1949, 1967). *Thermodynamics: An Advanced Treatment for Chemists and Physicists*. Amsterdam: North-Holland Publishing.
258. Habermas, J. (1985, 1987). *The Philosophical Discourse of Modernity*. Cambridge: MIT Press.
259. Hacking, I. (1975, 2009). *The emergence of probability: a philosophical study of early ideas about probability induction and statistical inference* New York: Cambridge University Press.
260. Haddad, W. M., Chellaboina, V., & Nersesov, S. G. (2005). *Thermodynamics: A Dynamical Systems Approach*. Princeton: Princeton University Press.
261. Hajek, A. (2012). Miracles. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*: <<http://plato.stanford.edu/archives/win2012/entries/probability-interpret/>>.
262. Hájek, A. (2007). Interpretations of probability. In S. University (Ed.), *Stanford Encyclopedia of Philosophy*: Internet.

263. Hamilton, A. G. (1982). *Numbers, sets and axioms. The apparatus of mathematics*. Cambridge: Cambridge University Press.
264. Hamilton, M. B. (2009). *Population Genetics*. Chichester: Wiley-Blackwell.
265. Hammond, G. P. (2004). "Engineering sustainability: thermodynamics, energy systems, and the environment". *International Journal of Energy Research*, 28, 613-639.
266. Handfield, T. (2012). *A philosophical guide to chance*. New York: Cambridge University Press.
267. Hasselblatt, B., & Katok, A. (2010). Handbook of Dynamical Systems - Volume 1. Survey I - Principal Structures. In: Internet 2010.
268. Hayward, S. A. (1999). "Relativistic thermodynamics". *arXiv-gr-qc: 9803007v2*.
269. Healey, R. (1994). "Nonseparable process and causal explanation". *Studies in History and Philosophy of Modern Physics*, 25, 337-374.
270. Heidegger, M. (1958). La question de la technique. In *Essais et conférences*. Paris: Gallimard.
271. Heil, J. (1999). "Multiple Realizability". *American Philosophical Quarterly*, 36, 189.
272. Hempel, C. (1965). *Philosophy of natural science*. New Jersey: Prentice Hall.
273. Hempel, C., & Oppenheim, P. (2008). On the Idea of Emergence. In M. A. Bedau & P. Humphreys (Eds.), *Emergence: Contemporary Readings in Philosophy and Science*. Cambridge: The MIT Press.
274. Henderson, L. (2003). "Von Neumann entropy: a reply to Shenker". *British Journal for the Philosophy of Science*, 54, 291-296.
275. Hendricks, V., & Symons, J. (2006). Epistemic logic. In S. University (Ed.), *Stanford Encyclopedia of Philosophy* (2 fév. 2003 ed.).
276. Hertz, J. (2004). "Historique en grandes enjambées de la thermodynamique de l'équilibre". *Journal de Physique*, 4, 3-20.
277. Heylighen, F. (2002). The Science of Self-Organization and Adaptivity. In *The Encyclopedia of Life Support Systems*.
278. Hilbert, D. (1926). "Sur l'infini". *Acta Mathematica*, 48, 91-122.
279. Hobbes, T. (1651, 1985). *Leviathan*. London: Penguin Classics.
280. Hoelzer, G. A., Smith, E., & Pepper, J. W. (2006). "On the logical relationship between natural selection and self- organization". *Journal of the European Society for Evolutionary Biology*, 19, 1785-1794.
281. Hoelzer, G. A., Smith, E., & Pepper, J. W. (2006). "On the logical relationship between natural selection and self- organization". *Journal compilation of the european society for evolutionary biology*, 19, 1785-1794.
282. Holton, G., & Brush, S. G. (2001, 2005). *Physics, the Human Adventure. From Copernicus to Einstein, and Beyond*. New Brunswick: Rutgers University Press.
283. Hooker, C. (2011). Philosophy of Complex Systems. In D. M. Gabbay, P. Thagard & J. Woods (Eds.), *Handbook of the Philosophy of Science* (Vol. 10). Oxford: Elsevier.
284. Horwich, P. (1987). *Asymmetries in time*. Cambridge: MIT Press.
285. Howard, I. K. (2001). "S Is for Entropy. U Is for Energy. What Was Clausius Thinking?". *Journal of Chemical Education*, 78, 505-508.
286. Hu, X., & Ye, Z. (2006). "Generalized quantum entropy". *Journal of Mathematical Physics*, 47.
287. Huby, P. M. (1971). "Kant or Cantor? That the Universe, If Real, Must Be Finite in Both Space and Time ". *Philosophy*, 46, 121-132.
288. Hume, D. (1739, 1969). *A Treatise of Human Nature*. London: Penguin Books.
289. Hume, D. (1739, 1995). *Traité de la nature humaine*. Paris: GF-Flammarion.
290. Hume, D. (1748, 1983). *Enquête sur l'entendement humain*. Paris: GF-Flammarion.
291. Humphreys, P. (1997). "Emergence, not supervenience". *Philosophy of Science*, 64 (Proceedings), 337-345.
292. Humphreys, P. (1997). "How Properties Emerge". *Philosophy of Science*, 64, 1-17.
293. Humphreys, P. (2008). "Computational and Conceptual Emergence". *Philosophy of Science*, 75, 584-594.
294. Huneman, P. (2008a). "Emergence made ontological? Computational versus combinatorial approaches". *Philosophy of Science*, 75, 595-607.
295. _____. (2008b). "Emergence and Adaptation". *Minds & Machines*, 18, 493-520.
296. _____. (2010a). "Determinism, predictability and open-ended evolution: lessons from computational emergence". *Synthese*, 11 March 2011.
297. _____. (2010b). "Assessing the Prospects for a Return of Organisms in Evolutionary Biology". *History and Philosophy of the Life Sciences*, 32, 341-372.

298. Huneman, P., & Humphreys, P. (2008). "Dynamical Emergence and Computation: An Introduction". *Minds & Machines*, 18.
299. Hutcheson, K. (1982). "What Happened to Occult Qualities in the Scientific Revolution?". *Isis*, 73, 233-253.
300. Hüttemann, A. (2005). "Explanation, emergence and quantum entanglement". *Philosophy of Science*, 72, 114-127.
301. Hüttemann, A., & Terzidis, O. (2000). "Emergence in physics". *International studies in the philosophy of science*, 14, 267-282.
302. Isham, C. J., & Savvidou, K. N. (2002). Time and Modern Physics. In K. Ridderbos (Ed.), *Time*. Cambridge: Cambridge University Press.
303. Ismael, J. T. (2009). "Probability in deterministic physics". *The Journal of Philosophy*, CVI, 89-108.
304. Jacob, F. (1970). *La logique du vivant. Une histoire du vivant*. Paris: Gallimard.
305. James, W. (1908). "The Meaning of the Word 'Truth' ". *Mind*, 17, 455-456.
306. Jammer, M. (1974). *The philosophy of quantum mechanics*. New York: John Wiley & Sons.
307. Jancel, R. (1963). *Les fondements de la mécanique statistique classique et quantique*. Paris: Gauthier-Villars.
308. Jaynes, E. T. (1965, 1983). Gibbs vs. Boltzmann Entropies. In R. D. Rosenkrantz (Ed.), *Papers on Probability, Statistics and Statistical Physics* (pp. 391-398). Dordrecht: Reidel.
309. Jaynes, E. T. (1995). Probability Theory: The Logic of Science. In Jaynes (Ed.): Internet.
310. Jodoin, L. (2009). Causality, quantum mechanics and entropy increase In C. c. d. s. humaines (Ed.). Ottawa: Société canadienne d'histoire et philosophie des sciences.
311. _____. (2010a). Histoire des sciences et des techniques modernes. In. École polytechnique de Montréal.
312. _____. (2010b). "Processus causal et intrication quantique". *Ithaque : revue de philosophie de l'Université de Montréal*, 6, 111-131.
313. _____. (2010c). Entropy and quantum measurement. In C. c. d. s. humaines (Ed.). Montréal: Société canadienne d'histoire et philosophie des sciences.
314. _____. (2010d). *Causalité et mécanique quantique*. Sarrebruck: Éditions universitaires européennes.
315. _____. (2010e). "L'héritage intellectuel de Mario Bunge: entre science et philosophie". *Philosophiques*, 37.
316. _____. (2011). "L'infini dans la théorie ergodique". *Ithaque : revue de philosophie de l'Université de Montréal*, 8.
317. _____. (2014a). "Mécanique statistique et génétique des populations : stratégies explicatives et analogies formelles". *Lato sensu*, 1.
318. _____. (2014b). "Knowledge as Inferred Truth". *Philosophical Writings*, 42.
319. Joule, J. P. (1884). *Scientific Papers*. Londres: The Physical Society of London.
320. Kaila, V. R. I., & Annala, A. (2008). "Natural selection for least action". *Proceedings of the Royal Society A*, 464, 3055-3070.
321. Kaplan, M. (1985). "It's not what you know that counts". *Journal of Philosophy*, 82, 350-363.
322. Kauffman, S., Logan, R. K., Este, R., Goebel, R., Hobill, D., & Shmulevich, I. (2007). "Propagating organization: an enquiry". *Biology and Philosophy*.
323. Kauffman, S. A. (1993). *The origin of order : self-organization and selection in evolution*. New York: Oxford University Press.
324. Kauffman, S. A. (1995). *At home in the universe : the search for laws of self-organization and complexity*. New York: Oxford University Press.
325. Keller, E. F. (200a). "Self-Organization, Self-Assembly, and the Origin of Life". *Boston Studies in the Philosophy of Science*, 266, 131-140.
326. Keller, E. F. (2011). "Towards a science of informed matter". *Studies in History and Philosophy of Biological and Biomedical Sciences*, 42, 174-179.
327. Khinchin, A. I. (1949). *Mathematical Foundations of Statistical Mechanics*. New York: Dover Publications inc.
328. Kim, J. (1999). "Making Sense of Emergence". *Philosophical Studies*, 95, 3-36.
329. Kitcher, P. (1981). "Explanatory unification". *Philosophy of Science*, 48.
330. Kleene, S. C. (1967). *Mathematical logic*. New York: John Wiley & Sons.
331. Klein, É. (1998). "Le temps de la physique". *Bulletin Interactif du Centre International de Recherches et Études transdisciplinaires*, 12.
332. Klein, É., & Sapiro, M. (1994). Le Temps et sa flèche. In. Paris: Flammarion.

333. Klein, M. J.. (1953). "Order, Organisation and Entropy". *British Journal for the Philosophy of Science*, 4, 158-160.
334. _____. (1967). "Thermodynamics in Einstein's Thought". *Science*, 157, 509-516.
335. Klein, M. (1970). "Maxwell, His Demon, and the Second Law of Thermodynamics". *American Scientist*, 58, 84-97.
336. _____. (1972). "Mechanical explanation at the end of the nineteenth century". *Centaurus*, 17.
337. _____. (1970). "Maxwell, His Demon, and the Second Law of Thermodynamics". *American Scientist*, 58, 84-97.
338. _____. (1972). "Mechanical explanation at the end of the nineteenth century". *Centaurus*, 17.
339. _____. (1974). "Carnot's contribution to thermodynamics". *Physics Today*, 27.
340. _____. (1983). The Scientific Style of Josiah Willard Gibbs. In R. Aris, H. T. Davis & R. H. Stuewer (Eds.), *Springs of Scientific Creativity. Essays on Founders of Modern Science*. Minneapolis: University of Minnesota Press.
341. Klir, G. J. (2006). *Uncertainty and information : foundations of generalized information theory*. New Jersey: John Wiley & Sons.
342. Kneebone, G. T. (1963, 2001). *Mathematical logic and the Foundations of Mathematics. An introductory survey*. New York: Dover Publications, inc.
343. Koenig, F. O. (1943). On the Significance of the Forgotten Thermodynamic Theorems of Carnot. In *Essays in Biology*: University of California Press.
344. Kojman-Rozen, S. (2011). "Time and Emergence in Victorian Scientific Theories: Lyell, Darwin and Maxwell". *Logic and Philosophy of Science*, IX, 285-293.
345. Kondepudi, D. (2008). *Introduction to Modern Thermodynamics*. Chichester: John Wiley & Sons.
346. Koons, R. C. (2004). "The logic of causal explanation: an axiomatization". *Studia Logica: An International Journal for Symbolic Logic*, 77, 325-354.
347. Korn, R. W. (2005). "The emergence principle in biological hierarchies". *Biology and Philosophy*, 20, 137-151.
348. Kostic, M. M. (2008). Sadi Carnot's Ingenious Reasoning of Ideal Heat Engine Reversible Cycles. In T. P. e. al. (Ed.), *International Conference on EESD* (pp. 159-166). Algarve, Portugal: IASME/WSEAS.
349. Koyré, A. (1966, 2003). *Études d'histoire de la pensée scientifique*. Paris: Gallimard.
350. Kripke, S. (1975). "Outline of a Theory of Truth". *Journal of Philosophy*, 72, 690-716.
351. Kronz, F. M., & Tichen, J. T. (2002). "Emergence and quantum mechanics". *Philosophy of Science*, 69, 324-347.
352. Kuhn, T. S. (1955a). "Carnot's Version of "Carnot's Cycle"". *American Journal of Physics*, 23, 91-95.
353. _____. (1955b). "La Mer's Version of "Carnot's Cycle"". *American Journal of Physics (Letters to the Editor)*, 23, 387-389.
354. _____. (1959). Energy Conservation as an Example of Simultaneous Discovery. In M. Clagett (Ed.), *Critical Problems in the History of Science*. Madison: University of Wisconsin Press.
355. _____. (1962, 2003). *La structure des révolutions scientifiques*. Paris: Champs-Flammarion.
356. Kuipers, T. A. F. (2007). General Philosophy of Science: Focal Issues. In. Amsterdam: Elsevier.
357. Kümmel, R. (2011). *The Second Law of Economics: Energy, Entropy, and the Origins of Wealth*. New York: Springer.
358. Kupiec, J.-J. (2005). "Un ordre humain, trop humain...". *Sciences et Avenir. Hors-série : L'énigme de l'émergence*, 143, 30-34.
359. Kvanvig, J. L. (1998). "Why Should Inquiring Minds Want to Know?: "Meno" Problems and Epistemological Axiology". *Monist*, 81, 426-451.
360. La Mer, V. K. (1954). "Some Current Misinterpretations of N. L. Sadi Carnot's Memoir and Cycle". *American Journal of Physics*, 22, 20-27.
361. La Mer, V. K. (1955). "Some Current Misinterpretations of N. L. Sadi Carnot's Memoir and Cycle II". *American Journal of Physics*, 23, 95-102.
362. Ladyman, J., Presnell, S., Short, A. J., & Groisman, B. (2007). "The connection between logical and thermodynamic irreversibility". *Studies in History and Philosophy of Modern Physics*, 38, 58-79.
363. Lakatos, I. (1978). *The methodology of scientific research programmes*. Cambridge: Cambridge University Press.
364. Lamy, J., & Gingras, Y. (2008). The Relationships Between Astronomical Observatories and Universities in nineteenth-century France. In M. Feingold (Ed.), *History of Universities* (Vol. 23). Oxford: Oxford University Press.

365. Landau, L. D., & Lifshitz, E. M. (1959, 1980). *Statistical Physics. Part I*. New York: Pergamon Press.
366. Landemore, H. (2004). *Hume, probabilité et choix raisonnable*. Paris.
367. Landsberg, P. T. (1985). "Usages et limites du concept d'entropie". *Communications*, 41, 63-80.
368. Lange, M. (2002). Philosophy of Science. An Anthology. In. Oxford: Blackwell Publishing Ltd.
369. _____. (2005). "Laws and their stability". *Synthese*, 144, 415-432.
370. _____. (2007). "Laws and meta-laws of nature : Conservation laws and symmetries". *Studies in History and Philosophy of Modern Physics*, 38, 457-481.
371. Lebowitz, J. L. (1999). "Statistical Mechanics: A Selective Review of Two Central Issues". *Reviews of Modern Physics*, 71, S346-S357.
372. Lebowitz, J. L. (1999). "Microscopic Origins of Irreversible Macroscopic Behavior". *Physica A*, 263, 516-527.
373. Leeds, S. (2003). "Foundations of Statistical Mechanics— Two Approaches". *Philosophy of Science*, 70, 126-144.
374. Leff, H. S. (2007). "Entropy, Its Language, and Interpretation". *Foundations of Physics*, 37, 1744-1766.
375. Lenk, H., & Stephan, A. (2002). On Levels and Types of Complexity and Emergence. In E. Agazzi & L. Montecucco (Eds.), *Complexity and Emergence*. Singapour: World Scientific Publishing Co.
376. Lervig, P. (1972). "On the Structure of Carnot's Theory of Heat". *Archives for the History of Exact Sciences*, 9, 222-239.
377. Lewis, D. (1996). "Elusive knowledge". *Australasian Journal of Philosophy*, 74, 549-567.
378. Lieb, E. H., & Yngvason, J. (1999). "The Physics and Mathematics of the Second Law of Thermodynamics". *Physics Report*, 310, 1-96.
379. _____. (2003). The Mathematical Structure of the Second Law of Thermodynamics. In *arXiv.org*: Internet.
380. Lindley, D. (2004). *Degrees Kelvin : a tale of genius, invention, and tragedy*. Washington: Joseph Henry Press.
381. Lindsay, R. B., & Margenau, H. (1936, 1957). *Foundations of Physics*. New York: Dover.
382. Lith, J. v. (2001). "Ergodic Theory, Interpretations of Probability and the Foundations of Statistical Mechanics". *Studies in History and Philosophy of Modern Physics*, 32, 581-594.
383. Liu, C. (2000). "Infinite Systems in SM Explanations: Thermodynamic Limit, Renormalization (semi-)Groups, and Irreversibility". *Philosophy of Science*, 68, S325-344.
384. Lloyd, E. A., & Anderson, C. G. (1993). Empiricism, objectivity, and explanation. In P. A. F. e. al. (Ed.), *Midwest Studies in Philosophy* (Vol. XVII). Notre Dame: Notre Dame University Press.
385. Loewer, B. (2007). Counterfactuals and the Second Law. In H. Price & R. Corry (Eds.), *Causation, Physics, and the Constitution of Reality*. New York: Clarendon Press.
386. Lubkin, E. (1987). "Keeping the Entropy of Measurement: Szilard Revisited". *International Journal of Theoretical Physics*, 26, 523-535.
387. Lucrèce. (1984). *De la nature*. Paris: Tel - Gallimard.
388. Ludovisi, A., Pandolfi, P., & Taticchi, M. I. (2005). "The strategy of ecosystem development: specific dissipation as an indicator of ecosystem maturity". *Journal of Theoretical Biology*, 235, 33-43.
389. Luminet, J.-P. (1994). Matière, espace, temps. In É. Klein & M. Sapiro (Eds.), *Le Temps et sa flèche*. Paris: Flammarion.
390. Lundgreen, P. (1990). "Engineering Education in Europe and the USA, 1750-1930". *Annals of Science*, 47, 33-75.
391. Lupasco, S. (1962). *L'énergie et la matière vivante*. Paris: Julliard.
392. Lyotard, J.-F. (1979). *La condition post-moderne*. Paris: Les Éditions de Minuit.
393. Lyre, H. (2008). "Time in Philosophy of Physics: The Central Issues". *Physics and Philosophy*.
394. Machamer, P., Darden, L., & Craver, C. F. (2000). "Thinking about mechanisms". *Philosophy of Science*, 67, 1-25.
395. Mackey, G. W. (1974). "Ergodic Theory and Its Significance for Statistical Mechanics and Probability Theory". *Advances in Mathematics*, 12, 178-268.
396. Mackey, M. C. (1992). *Time's Arrow: The origins of thermodynamic behavior*. New York: Springer-Verlag.
397. Mackie, J. L. (1974, 1980). *The cement of the universe*. Oxford: Oxford University Press.
398. Mader, S. S. (1988). *Biologie. Évolution, diversité et environnement*. Saint-Laurent: Éditions du Trécaré inc.
399. Magie, W. F. (1899). The second law of thermodynamics: Memoirs by Carnot, Clausius and Thomson. In J. S. Ames (Ed.), *Harper's scientific memoirs*. New York: Haper & Brothers Publishers.
400. Mainzer, K. (2005). *Symmetry and Complexity. The Spirit and Beauty of Nonlinear Science*. Singapore: World Scientific Publishin.

401. Mainzer, K. (2010). "The Emergence of Temporal Structures in Dynamical Systems". *Foundations of Physics*, 40, 1638-1650.
402. Malaterre, C. (2010). *Les origines de la vie : émergence ou explication réductive ?* Paris: Hermann.
403. Mantegna, R. N., & Stanley, H. E. (2000). *An Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge: Cambridge University Press.
404. Margolis, E., & Laurence, S. (2014). Concepts. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*: Internet.
405. Margulis, L., & Sagan, D. (2000, 2002). *What Is Life?* Berkeley: University of California Press.
406. Maroney, O. J. E. (2008). "The Physical Basis of the Gibbs-von Neumann entropy.". *arXiv:quant-ph/0701127v2*.
407. Martin, T. (2007). Le tout et les parties dans les systèmes naturels. In. Paris: Vuibert.
408. _____. (2009). L'unité des sciences : nouvelles perspectives. In. Paris: Vuibert.
409. Martyusheva, L. M., & Selezneva, V. D. (2006). "Maximum entropy production principle in physics, chemistry and biology". *Physics Reports*, 26, 1-45.
410. Matthen, M., & Arieu, A. (2002). "Two Ways of Thinking about Fitness and Natural Selection". *Journal of Philosophy*, 99, 55-83.
411. Matthen, M., & Stephens, C. (2007). Philosophy of Biology. In D. M. Gabbay, P. Thagard & J. Woods (Eds.), *Handbook of the Philosophy of Science*. Amsterdam: Elsevier.
412. Maxwell, G. (1962). "The Ontological Status of Theoretical Entities". *Minnesota Studies in Philosophy of Science*, III.
413. Maxwell, J. C. (1860). Illustrations of the Dynamical Theory of Gases. In S. G. Brush (Ed.), *The Kinetic Theory of Gases*. London: Imperial.
414. _____. (1867). On the Dynamical Theory of Gases. In S. G. Brush (Ed.), *The Kinetic Theory of Gases*. London: Imperial.
415. _____. (1871, 1902). *Theory of Heat*. London: Longmans, Green & co.
416. _____. (1878). "Tait's "Thermodynamics"". *Nature*, 17, 257-259.
417. Mayer, B., Köhler, G., & Rasmussen, S. (1997). "Simulation and dynamics of entropy-driven, molecular self-assembly processes". *Physical Review A*, 55, 4489-4499.
418. Mayr, E. (1997). "The objects of selection". *Proceedings of the National Academy of Sciences of the USA*, 94, 2091-2094.
419. Mayr, E. (2004, 2006). *Après Darwin : La biologie, une science pas comme les autres*. Paris: Dunod.
420. Mayumi, K. (2001). *The Origins of Ecological Economics : The bioeconomics of Georgescu-Roegen*. Londres: Routledge.
421. Mazauric, S. (2009). *Histoire des sciences à l'époque moderne*. Paris: Armand Colin.
422. McGrew, T. (2014). Miracles. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*: <http://plato.stanford.edu/archives/win2014/entries/miracles/>.
423. McLaughlin, B., & Bennett, K. (2011). Supervenience. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*: <http://plato.stanford.edu/archives/win2011/entries/supervenience>.
424. McTaggart, J. E. (1908). "The Unreality of Time". *Mind*, 17, 457-474.
425. Merleau-Ponty, J. (1979). "La découverte des principes de l'énergie : l'itinéraire de Joule". *Revue d'histoire des sciences*, 32, 315-331.
426. Monod, J. (1970). *Le hasard et la nécessité*. Paris: Éditions du Seuil.
427. Morowitz, H. J. (2002). *The Emergence of Everything: How the World Became Complex*. Oxford: Oxford University Press.
428. Morrison, M. (1997). "Physical Models and Biological Contexts". *Philosophy of Science*, 64, S315-324.
429. _____. (2004). "Population Genetics and Population Thinking: Mathematics and the Role of the Individual". *Philosophy of Science*, 71, 1189-1200.
430. _____. (2006a). "Unification, Explanation and Explaining Unity: The Fisher-Wright Controversy". *British Journal for the Philosophy of Science*, 57, 233-245.
431. _____. (2006b). "Emergence, Reduction, and Theoretical Principles: Rethinking Fundamentalism". *Philosophy of Science*, 73.
432. _____. (2007). The Development of Population Genetics. In M. Matthen & C. Stephens (Eds.), *Philosophy of Biology*. Amsterdam: Elsevier.
433. Müller, I. (2010). Entropy: A subtle concept in thermodynamics. In A. Greven, G. Keller & G. Warnecke (Eds.), *Entropy*. Princeton: Princeton University Press.

434. Müller, I. (2007). *A History of Thermodynamics. The Doctrine of Energy and Entropy*. Berlin, Heidelberg: Springer-Verlag.
435. Müller, I., & Weiss, W. (2005). *Entropy and Energy. A Universal Competition*. Heidelberg: Springer-Verlag.
436. _____. (2008). "Entropy and Energy – a Universal Competition". *Entropy*, 10, 462-476.
437. Murthy, K. P. N. (2006). "Ludwig Boltzmann, Transport Equation and the Second law". *arXiv:cond-mat/0601566v3 [cond-mat.stat-mech]*.
438. Musson, A. E., & Robinson, E. (1969). *Science and Technology in the Industrial Revolution*. Manchester: Manchester University Press.
439. Navarro, L. (1998). "Gibbs, Einstein and the Foundations of Statistical Mechanics". *Archives for the History of Exact Sciences*, 53, 147-180.
440. Neumaier, A. (2007). On the foundations of thermodynamics. In *arXiv:0705.3790v1*: Internet.
441. Neumann, J. v. (1955). *The Mathematical Foundations of Quantum Mechanics*. Princeton: Princeton University Press.
442. Newburgh, R. (2009). "Carnot to Clausius: caloric to entropy". *European Journal of Physics*, 30, 713-728.
443. Newman, S. A., & Forgacs, G. (2005). Complexity and Self-organization in Biological Development and Evolution. In.
444. Nicolis, G. (2005). "Le mécanisme "bottom-up"". *Sciences et Avenir, Hors-série*.
445. Nicolis, G., & Prigogine, I. (1977). *Self-organization in non-equilibrium systems*. New York: Wiley-Interscience.
446. North, J. (2011). Time in Thermodynamics. In C. Callender (Ed.), *The Oxford Handbook of Philosophy of Time* (pp. 312-350). Oxford: Oxford University Press.
447. O'Connor, T., & Wong, H. Y. (2012). Emergent properties. In *Stanford Encyclopedia of Philosophy*: Internet.
448. Okasha, S. (2008). "Fisher's Fundamental Theorem of Natural Selection — A Philosophical Analysis". *British Journal for the Philosophy of Science*, 59, 319-351.
449. Osler, M. J. (2000). Early-modern protestantism. In G. B. Ferngren (Ed.), *The history of science and religion in the western tradition : an encyclopedia*. New York: Garland.
450. Pais, A. (1982). *Subtle is the Lord... The Science and life of Albert Einstein*. Oxford: Oxford University Press.
451. Parker, D. (2005). "Thermodynamic Irreversibility: Does the Big Bang Explain What It Purports to Explain?". *Philosophy of Science*, 72, 751-763.
452. Partovi, M. H. (2007). "Verschränkung versus Stosszahlansatz : Disappearance of the Thermodynamic Arrow in a High-Correlation Environment".
453. Pearl, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge: Cambridge University Press.
454. Pegg, D. T. (2006). "Causality in quantum mechanics". *Physics Letters A*, 349, 411-414.
455. Pegg, D. T. (2008). "Retrocausality and quantum mechanics". *Studies in History and Philosophy of Modern Physics*, 39, 830-840.
456. Penrose, O. (1970). *Foundations of statistical mechanics. A deductive treatment*. Oxford: Pergamon Press.
457. Penrose, O., & Percival, I. C. (1962). "The Direction of Time". *Proceedings of the Physical Society*, 79, 605-616.
458. Penrose, R. (2004). *The Road to Reality. A Complete Guide to the Laws of the Universe*. London: Jonathan Cape.
459. Peres, A. (2002). *Quantum Theory: Concepts and Methods*. New York: Kluwer Academic Publishers.
460. Planck, M. (1933, 1963). *L'image du monde dans la physique moderne*. Genève: Gonthier.
461. Planck, M. (1945). *Treatise on Thermodynamics*. New York: Dover Publications.
462. Poincaré, H. (1902, 1968). *La science et l'hypothèse*. Paris: Champs-Flammarion.
463. _____. (1908). *Thermodynamique. Cours de physique théorique*. Paris: Gauthier-Villard.
464. Popper, K. R. (1934, 1973). *La logique de la découverte scientifique*. Paris: Bibliothèque scientifique Payot.
465. _____. (1959). "The Propensity Interpretation of Probability". *The British Journal for the Philosophy of Science*, 10, 25-42.
466. _____. (1962, 1979). *La société ouverte et ses ennemis. Tome 1 : L'ascendant de Platon*. Paris: Seuil.
467. _____. (1962, 1979). *La société ouverte et ses ennemis. Tome 2 : Hegel et Marx*. Paris: Seuil.
468. _____. (1963, 2006). *Conjectures and refutations*. Oxon: Routledge.
469. _____. (1979, 1991). *La connaissance objective*. Paris: Champs-Flammarion.
470. Porter, R. (2008). The Cambridge History of Science. Volume 4 : The Eighteen Century. In. Cambridge: Cambridge University Press.

471. Pottier, N. (2007). *Physique statistique hors d'équilibre. Processus irréversibles linéaires*. Paris: EDP Sciences/CNRS Éditions.
472. Price, H. (1996). *Time's arrow and Archimedes' point*. Oxford: Oxford University Press.
473. Price, H., & Corry, R. (2007). Causation, Physics, and the Constitution of Reality. In. New York: Clarendon Press.
474. Prigogine, I. (1968, 1996). *Introduction à la thermodynamique des processus irréversibles*. Paris: Éditions Jacques Gabay.
475. Prigogine, I., Nicolis, G., & Babloyantz, A. (1972). "Thermodynamics of Evolution". *Physics Today*, 25.
476. Prigogine, I., & Stengers, I. (1979, 1986). *La nouvelle alliance*. Paris: Gallimard.
477. Provine, W. B. (1971, 2001). *The origins of theoretical population genetics : with a new afterword*. Chicago: University of Chicago Press.
478. Psillos, S. (2007). Past and Contemporary Perspectives on Explanation. In T. A. F. Kuipers (Ed.), *General Philosophy of Science: Focal Issues*. Amsterdam: Elsevier.
479. Psillos, S., & Curd, M. (2008). The Routledge Companion to Philosophy of Science. In. Abington: Routledge.
480. Quine, W. v. O. (1960). *Word and object*. Cambridge: M.I.T. Press.
481. _____. (1992). *Pursuit of truth*. Cambridge: Harvard University Press.
482. Rae, A. I. M. (2002). *Quantum Mechanics*. Bristol: Institute of Physics Publishing.
483. Rankine, W. J. M. (1865). "On the second law of the thermodynamics". *Philosophical Magazine*, 4, 241-245.
484. Rankine, W. J. M. (1867). "Sur la nécessité de vulgariser la seconde loi de la thermodynamique". *Annales de chimie et de physique*, 4.
485. Rawls, J. (1971, 2005). *A Theory of Justice*. Cambridge: Harvard University Press.
486. Redondi, P. (1976). "Sadi Carnot et la recherche technologique en France de 1825 à 1850". *Revue d'histoire des sciences*, 29, 243-259.
487. Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Jackson, R. B. (2005, 2011). *Biology* (9e ed.). San Francisco: Pearson Benjamin Cummings.
488. Regt, H. C. D. G. d. (2006). "To believe in belief: Popper and van Fraassen on Scientific Realism". *Journal for General Philosophy of Science*, 37, 21-39.
489. Reichenbach, H. (1956). *The direction of time*. Berkeley et Los Angeles: University of California Press.
490. Reif, F. (1965). *Fundamentals of Statistical and Thermal Physics*. New York: McGraw-Hill.
491. Reill, P. H. (2008). The Legacy of the "Scientific Revolution": Science and the Enlightenment. In R. Porter (Ed.), *The Cambridge History of Science. Volume 4 : The Eighteenth Century*. Cambridge: Cambridge University Press.
492. Reisman, K., & Forber, P. (2005). "Manipulation and the Causes of Evolution". *Philosophy of Science*, 72, 1113-1123.
493. Rényi, A. (1970, 2007). *Foundations of Probability*. Mineola: Dover Publications.
494. Resher, N. (1978, 1993). *Le progrès scientifique*. Paris: Presses universitaires de France.
495. _____. (2005). *Epistemic logic*. Pittsburgh: Pittsburgh University Press.
496. Revel, J.-F. (1968). *Histoire de la philosophie occidentale. Penseurs grecs et latins*. Paris: Stock.
497. Rhodes, C. J., & Demetrius, L. (2010). "Evolutionary Entropy Determines Invasion Success in Emergent Epidemics". *PLoS ONE*, 5, e12951.
498. Richmond, P., Mimkes, J., & Hutzler, S. (2013). *Econophysics and Physical Economics*. Oxford: Oxford University Press.
499. Rickles, D. (2011). Philosophy of Complex Systems. In C. Hooker (Ed.), *Handbook of the Philosophy of Science* (Vol. 10, pp. 531-565). Oxford: Elsevier.
500. Ridderbos, K. (2002). "The coarse-graining approach to statistical mechanics: how blissful is our ignorance?". *Studies in History and Philosophy of Modern Physics*, 33, 65-77.
501. Ridderbos, K. (2002). *Time*. Cambridge: Cambridge University Press.
502. Rigden, J. S. (2005, 2006). *Einstein 1905 : the standard of greatness*. Cambridge: Harvard University Press.
503. Rosenberg, A. (2000, 2005). *Philosophy of Science: A Contemporary Introduction*. New York: Routledge.
504. _____. (2001). "How is Biological Explanation Possible?". *The British Journal for the Philosophy of Science*, 52, 735-760.
505. _____. (2011). *The atheist's guide to reality* (Vol. 52). New York: W. W. Norton & Company.

506. Ross, D., & Spurrett, D. (2007). "Notions of cause: Russell's thesis revisited". *British Journal for the Philosophy of Science*, 1-32.
507. Rossi, P. (1999). *Aux origines de la science moderne*. Paris: Éditions du Seuil.
508. Rueger, A. (2000). "Physical Emergence, Diachronic and Synchronic". *Synthese*, 124.
509. Rüegg, W. (2004). A History of the University in Europe, Volume III. In. Cambridge: Cambridge University Press.
510. Rukeyser, M. (1942). *Willard Gibbs*. New York: Doubleday, Doran & Company, inc.
511. Russell, B. (1914, 2006). *Our knowledge of the external world*. Oxon: Routledge.
512. _____. (1948, 2005). *Human knowledge: Its scope and limits*. Oxon: Routledge.
513. Ruvinsky, A. (2010). *Genetics and randomness*. Boca Raton: Taylor & Francis Group.
514. Salem, W. K. A., & Fröhlich, J. (2006). "Status of the Fundamental Laws of Thermodynamics". *Journal of Statistical Physics*, 126, 1045-1068.
515. Salmon, W. (1978). "Why ask 'Why?'? An inquiry concerning scientific explanation". *Proceedings and Adresses of the American Philosophical Association*, 51, 683-705.
516. _____. (1984). "Scientific explanation: three basic conceptions". *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2, 293-305.
517. _____. (2005). *Reality and rationality*. New York: Oxford University Press.
518. Sarkar, S. (2007). Haldane and the Emergence of Modern Evolutionary Theory. In M. Matthen & C. Stephens (Eds.), *Philosophy of Biology*. Amsterdam: Elsevier.
519. Sarton, G., Mayer, J. R., Joule, J. P., & Carnot, S. (1929). "The Discovery of the Law of Conservation of Energy". *Isis*, 13, 18-44.
520. Satterthwaite, D. (1997). "Environmental Transformations in Cities as They Get Larger, Wealthier and Better Managed". *The Geographical Journal*, 163, 216-224.
521. Sawyer, R. K. (2004). "The Mechanisms of Emergence". *Philosophy of the Social Sciences*, 34, 260-282.
522. Scarani, V. (2006). *Initiation à la physique quantique*. Paris: Vuibert.
523. Schaffer, J. (2003). The metaphysics of causation. In S. University (Ed.), *Stanford Encyclopedia of Philosophy* (2 fév. 2003 ed.).
524. Schaffer, J. (2006). "Le trou noir de la causalité". *Philosophie*, 78-93.
525. Schickore, J. (2014). Scientific Discovery. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*: <http://plato.stanford.edu/archives/spr2014/entries/scientific-discovery/>.
526. Schmitt, C. (1922, 1988). *Théologie politique I*. Paris: Gallimard.
527. Schneider, E. D., & Kay, J. J. (1994). "Life as a Manifestation of the Second Law of Thermodynamics". *Mathematical and Computer Modelling*, 19, 25-48.
528. Schrödinger, E. (1944, 2010). *What is Life? with Mind and Matter and Autobiographical Sketches*. New York: Cambridge University Press.
529. Schulman, L. S., & Gaveau, B. (2001). "Coarse Grains: The Emergence of Space and Order". *Foundations of Physics*, 31.
530. Schwartz, J. (2008). *In pursuit of the gene: from Darwin to DNA*. Cambridge: Harvard University Press.
531. Seidenfeld, T. (1986). "Entropy and Uncertainty". *Philosophy of Science*, 53, 467-491.
532. Seidengart, J. (1999, 2003). Infini. In D. Lecourt (Ed.), *Dictionnaire d'histoire et philosophie des sciences*. Paris: Qadriga-PUF.
533. Sella, G., & Hirsh, A. E. (2005). "The application of statistical physics to evolutionary biology". *Proceedings of the National Academy of Sciences (USA)*, 102, 9541-9546.
534. Serrin, J. (1979). "Conceptual Analysis of the Classical Second Laws of Thermodynamics". *Archive for Rational Mechanics and Analysis*, 70.
535. Sesták, J., Mares, J. J., Hubík, P., & Proks, I. (2009). "Contribution by Lazara Carnot and Sadi Carnot to the caloric theory of heat and its inspirative role in thermodynamics". *Journal of Thermal Analysis and Calorimetry*.
536. Seth, S. (2007). "Crisis and the construction of modern theoretical physics". *British Journal for the History of Science*, 40, 25-51.
537. Sethna, J. P. (2011). *Statistical Mechanics : Entropy, Order Parameters, and Complexity*. Oxford: Clarendon Press.
538. Settle, T. (1975). Presuppositions of propensity theories of probability. In G. Maxwell & R. M. A. Jr (Eds.), *Induction, Probability, and Confirmation*. Minneapolis: University of Minnesota Press.
539. Shakespeare, W. (1994). *Complete works of William Shakespeare*. Glasgow: HarperCollins.

540. Shannon, C. E. (1948). "A Mathematical Theory of Communication". *The Bell System Technical Journal*, 27, 379-423; 623-356.
541. Shapin, S. (2008). *The scientific life*. Chicago: The University of Chicago Press.
542. Sharma, V., & Annala, A. (2007). "Natural process – Natural selection". *Biophysical Chemistry*, 127, 123-128.
543. Sheehan, D. P. (2007). "Thermosynthetic Life". *Foundations of Physics*, 37, 1774-1797.
544. Shenker, O. R. (1999). "Is -kTr(plnp) the Entropy in Quantum Mechanics". *The British Journal for the Philosophy of Science*, 50, 33-48.
545. Sieniutycz, S. (2000). "Some Thermodynamic Aspects of Development and Bistability in Complex Multistage Systems". *Open Systems & Information Dynamics*, 7, 309-326.
546. Silberstein, M., & McGeever, J. (1999). "The Search for Ontological Emergence". *The Philosophical Quarterly*, 49, 182-200.
547. Simon, H. A. (1962). "The Architecture of Complexity". *Proceedings of the American Philosophical Society*, 106, 467-482.
548. Skipper, R. (2007). Sir Ronald Aylmer Fisher. In M. Matthen & C. Stephens (Eds.), *Philosophy of Biology*. Amsterdam: Elsevier.
549. Sklar, L. (1973). "Statistical Explanation and Ergodic Theory". *Philosophy of Science*, 40, 194-212.
550. _____. (1993). *Physics and Chance*. Cambridge: Cambridge University Press.
551. _____. (1999). "The Reduction(?) of Thermodynamics to Statistical Mechanics". *Philosophical Studies*, 95, 187-202.
552. _____. (2013). *Philosophy and the foundations of dynamics*. Cambridge: Cambridge University Press.
553. Smith, C. (1976). "Natural Philosophy and Thermodynamics: William Thomson and 'The Dynamical Theory of Heat'". *The British Journal for the History of Science*, 9, 293-319.
554. Snow, C. P. (1959, 1961). *The two cultures and the scientific revolution*. Cambridge: Cambridge University Press.
555. Sober, E. (1983). "Equilibrium explanation". *Philosophical Studies*, 43, 201-210.
556. _____. (1984, 1985). *The Nature of Selection. Evolutionary Theory in Philosophical Focus*. Cambridge: MIT Press.
557. _____. (1993). *Philosophy of Biology*. Boulder: Westview Press.
558. _____. (1999). "The Multiple Realizability Argument Against Reductionism". *Philosophy of Science*, 66, 542-564.
559. Sosa, E. (2002). The Place of Truth in Epistemology. In M. DePaul & L. Zagzebski (Eds.), *Intellectual virtue : perspectives from ethics and epistemology*. Oxford: Oxford University Press.
560. Spire, A. (2007). "Énergie : sens et contresens". *Nouvelle Fondation*, 2, 40-44.
561. Stalnaker, R. (2006). "Of logics of knowledge and belief". *Philosophical Studies*, 128, 169-199.
562. Standish, R. K. (2001). "On Complexity and Emergence". *Complexity*, 9.
563. Stefanov, A., Zbinden, H., Gisin, N., & Suarez, A. (2002). "Quantum Correlations with Spacelike Separated Beam Splitters in Motion : Experimental Test of Multisimultaneity". *Physical Review Letters*, 88, 1-4.
564. Stewart, I. (2007). *Why beauty is truth : a history of symmetry*. New York: Basic Books.
565. Strandh, S. (1973). *Machines - Histoire illustrée*. Paris: Draeger.
566. Sturtevant, A. H. (1965, 2001). *A History of Genetics*. New York: Cold Spring Harbor Laboratory Press.
567. Suntsov, V. V., & Suntsova, N. I. (2008). "Concepts of macro- and microevolution as related to the problem of origin and global expansion of the plague Pathogen *Yersinia pestis*". *Biology Bulletin*, 35, 333-338.
568. Tarski, A. (1944). "The semantic conception of truth and the foundation of semantics". *Philosophy and Phenomenological Research*, 4, 341-376.
569. Taylor, C. (1992). *Grandeur et misère de la modernité*. Montréal: Bellarmin.
570. Teller, P. (1986). "Relational Holism and Quantum Mechanics". *The British Journal for the Philosophy of Science*, 37, 71-81.
571. Templeton, A. R. (2006). *Population genetics and microevolutionary theory*. New Jersey: John Wiley & Sons.
572. Thalos, M. (2002). "Explanation is a genus: an essay on the varieties of scientific explanation". *Syntheses*, 130, 317-354.
573. Thess, A. (2011). *The Entropy Principle. Thermodynamics for the Unsatisfied*. Berlin: Springer-Verlag.
574. Thomson, D. A. W. (1942). *On growth and form*. Cambridge: Cambridge University Press.

575. Thomson, W. (1849). "An Account of Carnot's Theory of the Motive Power of Heat; with Numerical Results deduced from Regnault's Experiments on Steam". *Transactions of the Edinburgh Royal Society*, 16.
576. _____. (1882). *Mathematical and Physical Papers. Volume I*. Cambridge: Cambridge University Press.
577. _____. (1882). *Mathematical and Physical Papers. Volume II*. Cambridge: Cambridge University Press.
578. Thorén, H., & Persson, J. (2013). "The Philosophy of Interdisciplinarity: Sustainability Science and Problem-Feeding". *Journal for General Philosophy of Science*, 44, 337-355.
579. Toda, M., Kubo, R., & Saitô, N. (1983, 1998). *Statistical Physics I. Equilibrium statistical mechanics*. Berlin: Springer.
580. Tolman, R. C. (1938, 1950). *The Principles of Statistical Mechanics*. London: Oxford University Press.
581. Tribus, M., & McIrvine, E. C. (1971). "Energy and information". *Scientific American*, 224, 178-184.
582. Truesdell, C. (1980). *The Tragicomical History of Thermodynamics, 1822-1854*. New York: Springer-Verlag.
583. Truesdell, C., & Bharatha, S. (1977). *The Concepts and Logic of Classical Thermodynamics as a Theory of Heat Engines*. New York: Springer-Verlag.
584. Tsallis, C., & Brigatti, E. (2004). "Nonextensive statistical mechanics: A brief introduction". *Continuum Mechanics and Thermodynamics*, 16, 223-235.
585. Tuckerman, M. E. (2010). *Statistical Mechanics: Theory and Molecular Simulation*. Oxford: Oxford University Press.
586. Turner, J. S. (2000). *The extended organism : the physiology of animal-built structures*. Cambridge/Londres: Harvard University Press.
587. Uffink, J. (2006). Compendium of the foundations of classical statistical physics. In: Internet.
588. _____. (2007). Compendium of the foundations of classical statistical physics. In J. Butterfield & J. Earman (Eds.), *Philosophy of Physics. Part B*. Amsterdam: Elsevier.
589. Uffink, J. (1995). "Can the Maximum Entropy Principle be Explained as a Consistency Requirement?". *Studies in History and Philosophy of Modern Physics*, 26, 223-261.
590. Uffink, J. (2001). "Bluff Your Way in the Second Law of Thermodynamics". *Studies in History and Philosophy of Modern Physics*, 32, 305-394.
591. Uspensky, V. A. (1992). "Kolmogorov and Mathematical Logic". *Journal of Symbolic Logic*, 57, 385-412.
592. van Fraassen, B. C. (1980). *The Scientific Image*. Oxford: Clarendon Press.
593. _____. (2008). *Scientific Representation: Paradoxes of Perspective*. Oxford: Clarendon Press.
594. Vassails, G. r. (1950). "Le poids du feu". *Revue d'histoire des sciences et de leurs applications*, 3, 222-241.
595. Vattimo, G. (2011). *A Farewell to Truth*. New York: Columbia University Press.
596. Vivien, F.-D. (2005). *Le développement soutenable*. Paris: La Découverte.
597. Vladar, H. P. d., & Barton, N. H. (2011). "The contribution of statistical physics to evolutionary biology". *arXiv:1104.2854v1, [q-bio.PE]*.
598. von Mises, R. (1928, 1957). *Probability, statistics, and truth*. New York: Dover.
599. Walsh, E. J., Davies, M. R. D., & McEligot, D. M. (2004). "On the use of Entropy to Predict Boundary Layer Stability". *Entropy*, 6, 375-387.
600. Walter, S. (2006). "Multiple Realizability and Reduction: A Defense of the Disjunctive Move". *Metaphysica*, 7, 43-65.
601. Walter T. Grandy, J. (2008). *Entropy and the Time Evolution of Macroscopic Systems*. Oxford: Oxford University Press.
602. Watson, P. (2005). *Ideas : a history from fire to Freud*. Londres: Orion.
603. Weber, B. H., & Depew, D. J. (1996). "Natural Selection and Self-Organization". *Biology and Philosophy*, 11, 33-65.
604. Weber, B. H., Depew, D. J., Dyke, C., Salthe, S. N., Schneider, E. D., Ulanowicz, R. E., & Wicken, J. S. (1989). "Evolution in Thermodynamic Perspective: An Ecological Approach". *Biology and Philosophy*, 4, 373-405.
605. Wehrl, A. (1978). "General properties of entropy". *Review of Modern Physics*, 50.
606. _____. (1991). "The Many Facets of Entropy". *Reports on Mathematical Physics*, 30, 119-129.
607. Weightman, G. (2007). *The industrial Revolutionaries*. Londres: Atlantic Books.
608. Weinberg, J. M., Nichols, S., & Stich, S. (2001). "Normativity and Epistemic Intuitions". *Philosophical Topics*, 29, 429-460.
609. Werndl, C. (2009). "Are deterministic descriptions and indeterministic descriptions observationally equivalent?". *Studies in History and Philosophy of Modern Physics*, 40, 232-242.

610. Wilkins, J. S., Stanyon, C., & Musgrave, I. (2012). "Selection without replicators: the origin of genes, and the replicator/interactor distinction in etiology". *Biology and Philosophy*, 27, 215-239.
611. Williams, P. E. (2002). "Energy and Entropy as the Fundaments of Theoretical Physics". *Entropy*, 4, 128-141.
612. Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.
613. Willigenburg, L. G. V., & Koning, W. L. D. (2009). "Emergence of the Second Law out of Reversible Dynamics". *Foundations of Physics*, 39, 1217-1239.
614. Wimsatt, W. (1981, 2012). Robustness, Reliability, and Overdetermination. In L. S. e. al. (Ed.), *Characterizing the Robustness of Science* (Vol. 292, pp. 61-87). Boston: Springer.
615. _____. (1994). "The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets". *Canadian Journal of Philosophy; Supplementary Volume*, 20, 207-274.
616. _____. (2000). "Emergence as non-aggregativity and the biases of reductionisms". *Foundations of Science*, 5, 267-296.
617. _____. (2006). "Aggregate, composed, and evolved systems: Reductionistic heuristics as means to more holistic theories". *Biology and Philosophy*, 21, 667-702.
618. _____. (2008). Aggregativity: Reductive Heuristics for Finding Emergence. In M. A. Bedau & P. Humphreys (Eds.), *Emergence: Contemporary Readings in Philosophy and Science* (pp. 99-110). Cambridge: The MIT Press.
619. Witmer, D. G. (2006). Supervenience. In S. Sarkar & J. Pfeifer (Eds.), *The philosophy of science : an encyclopedia*. New York: Routledge.
620. Wittgenstein, L. (1922, 1993). *Tractatus logico-philosophicus*. Paris: Gallimard.
621. Wolpert, D. H. (2008). "Physical limits of inference". *Physica D*, 237, 1257-1281.
622. Woodward, J. (1979). "Scientific explanation". *The British Journal for the Philosophy of Science*, 30, 41-67.
623. _____. (2003a). Scientific explanation. In S. University (Ed.), *Stanford Encyclopedia of Philosophy*: Internet.
624. _____. (2003b). *Making things happen. A theory of causal explanation*. New York: Oxford University Press.
625. Workman, R. W. (1964). "What makes an explanation". *Philosophy of Science*, 31, 241-254.
626. Wrigley, E. A. (2010). *Energy and the English Industrial Revolution*. Cambridge: Cambridge University Press.
627. Yolton, J. W. (1959). "Explanation". *The British Journal for the Philosophy of Science*, 10, 194-208.
628. Yoshioka, D. (2007). *Statistical Physics. An introduction*. Berlin: Springer-Verlag.
629. Yourgau, W., Merwe, A. v. d., & Raw, G. (1966). *Treatise on Irreversible and Statistical Thermodynamics: an Introduction to Nonclassical Thermodynamics*. New York: Macmillan.
630. Zagzebski, L. (1994). "The inescapability of the Gettier problem". *The Philosophical Quarterly*, 44, 65-73.
631. Zanchini, E., & Beretta, G. P. (2008). Rigorous Axiomatic Definition of Entropy Valid Also for Non-Equilibrium States. In G. P. Beretta, A. F. Ghoniem & G. N. Hatsopoulos (Eds.), *Meeting the Entropy Challenge, An International Thermodynamics Symposium*: American Institute of Physics.
632. Zeh, H.-D. (1989, 1992). *The Physical Basis of the Direction of Time* (Seconde ed.). Heidelberg: Springer-Verlag.
633. Zurek, W. H. (1989). "Algorithmic Randomness and Physical Entropy". *Physical Review A*, 40, 4731-4751.

9 Présentation de l’auteur

Après avoir complété des études en sciences appliquées puis en génie mécanique, j’ai bifurqué vers la philosophie avec un mémoire de maîtrise « Causalité et mécanique quantique » sous la direction de Michel Bitbol (Paris) et Marie-Hélène Parizeau (Québec).

J’ai enseigné à Polytechnique Montréal des cours d’histoire des sciences modernes et d’éthique appliquée, un cours au département de philosophie à l’université du Québec à Trois-Rivières ainsi qu’un séminaire de deuxième cycle à l’université de Montréal.

